

Pricing Tools and Analysis for Emerging e-Commerce Technologies

Michael Levi Hamilton

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Pricing Analysis and Tools for Emerging e-Commerce

Michael Levi Hamilton

With the deluge of big data, many retailers are experimenting with rich, data-driven pricing strategies. In this dissertation we study three emerging pricing strategies: (i) Opaque pricing, the pricing of products where some feature is hidden from the customer until after purchase. In a general model we give a sharp characterization for when opaque selling outperforms traditional forms of differentiated pricing. (ii) Personalized pricing, i.e. pricing strategies that predict an individual customer's valuation for a product and then offers them a customized price. Leveraging natural statistics of the valuation distribution, we prove tight upper and lowers on the ratio between personalized pricing strategies and simpler selling strategies, which, among other things, yields insight into which markets personalized pricing is most valuable. (iii) Loot box pricing, the pricing of (random) bundles of virtual items, the contents of which are revealed after purchase. In an asymptotic regime we compare and contrast the revenue of different forms of loot box pricing with traditional selling models, and give theory to explain the recent proliferation of loot boxes in mobile gaming markets.

Contents

List of Figures	iv
List of Tables	ix
Acknowledgements	x
Introduction	1
0.1 Introduction to Pricing for Revenue Maximization	2
0.2 Pricing a Single Good	3
0.3 Pricing Beyond A Fixed Market and Single Good	4
0.4 Summary of Contributions	6
Chapter 1 Opaque Selling	8
1.1 Introduction	9
1.2 Selling Models	16
1.3 The Power of Opaque Products	23
1.4 The Power of One Opaque Product	36
1.5 Numerical Experiments	38
1.6 Conclusion	42
Chapter 2 Personalized Pricing	44
2.1 Introduction	45
2.2 Model and Preliminaries	54

2.3	The Value of Personalized Pricing over Single Pricing	57
2.4	From Third-Degree to First-Degree Price Discrimination	71
2.5	Extensions	80
2.6	Conclusions	89
Chapter 3 Loot Box Selling		91
3.1	Introduction	92
3.2	Model and Preliminaries	97
3.3	Asymptotic Analysis of Loot Box Pricing	105
3.4	Loot Box Design Problem	109
3.5	Numerical Tests	119
3.6	Conclusions	122
Bibliography		124
Appendices		132
Chapter A Opaque Selling		133
A.1	A Primer on Discriminatory Pricing	133
A.2	An Extension of Exchangeability	134
A.3	Missing Examples from Chapter 1	137
A.4	Additional Proofs from Chapter 1	140
A.5	Auxiliary Lemmas	152
Chapter B Personalized Pricing		157
B.1	Bounding the Value of Personalized Pricing Using the General Moments	157
B.2	Omitted Proofs from Chapter 2	160
B.3	A Dynamic Programming Algorithm for Computing (kP)	189
Chapter C Loot Box Selling		191

C.1 Omitted Proofs from Chapter 3	191
---	-----

List of Figures

1.1	Swingline offers their “SmartTouch” staplers on Amazon.com traditionally alongside a “colors may vary” option i.e., a single opaque product where the color of the stapler is hidden until after purchase. The opaque product (right) is offered at the discounted price of \$15.99, which is \$1.63 less than the traditional price (left) of \$18.62.	9
1.2	SwimOutlet.com offers various styles of Nike brand swim trunks for prices between \$43.00-\$50.00 alongside an opaque “Grab Bag” for \$23.00. . . .	10
1.3	Eurowings.com uses an opaque selling strategy to offer airline tickets with a base price of €66.00. Here the destination of the flight is opaque, with $N = 12$ possible destinations. The site allows the customer to exclude as many destinations as they desire, each for a price of €5.00. In the figure above, three destinations are excluded and the price for the desired opaque product is €81.	12
1.4	Above are two valuation spaces for opaque selling strategies with prices $(p, p^2) = (4, 3)$. Left is the valuation space for pessimistic customers, right is valuation space for risk-neutral customers. The darkened regions correspond to customer valuations that yield purchases of an item at a price of 4. The lighter regions correspond to purchases of the opaque product which has a price of 3. The unshaded regions correspond to valuations that yield no purchase.	19

1.5	The valuation space and purchasing behaviors for a pessimistic customer facing OPQ and DP selling strategies respectively. Customers with valuations in the darkened regions buy at price 5 in both figures. Customers with valuations in the lightly shaded regions buy at price 3 (i.e., purchase the opaque product or item 2, respectively). Customers in the unshaded region do not purchase.	28
1.6	Illustrates the relationship between \mathcal{R}_{SP} (dashed line), \mathcal{R}_{DP} (dotted line), and \mathcal{R}_{OPQ} (solid line) as the proportion of pessimistic customers increases.	41
1.7	Illustrates the relationship between \mathcal{R}_{SP} , \mathcal{R}_{DP} , \mathcal{R}_{1OPQ} and \mathcal{R}_{OPQ} as the proportion of pessimistic customers increases.	42
2.1	We represent our pricing models in terms of their <i>price flexibility</i> (y-axis) and <i>prediction accuracy</i> (x-axis).	50
2.2	The left panel shows the two real branches of the Lambert-W function, $W_0(\cdot)$ (solid black), and $W_{-1}(\cdot)$ (dashed). Our bounds depend upon the $W_{-1}(\cdot)$ branch (rescaled), as shown in right panel, and which can be upper and lower bounded via Chatzigeorgiou [29] (dotted).	56
2.3	The left panel shows the tight distribution of Lemma 2.3.2 when $M = 1$ and $S = 5$. The middle panel shows the bound in Lemma 2.3.2 when $M = 1$ and as S varies between 1 and 10. The right panel shows the bound in Lemma 2.3.2 when $S = 5$ and as M varies between 0.1 and 1.0.	61
2.4	The left panel plots the bound from Theorem 2.3.1 as a function of D with $S = 4$ and $M = .9$. The right panel plots the inverse of this bound, which we note is convex.	65
2.5	Tight distributions for Theorem 2.3.1 in each regime when $S = 4$, $\mu = 1$ and $M = 1$	67

2.6	The left panel shows the revenue of a single price strategy (dark colored square) in relation to the supporting lower bound (light colored trapezoid) for a unimodal cCDF. The right panel plots the guarantee of Theorem 2.3.2 (green) against the upper bound of Theorem 2.3.1 (red).	70
2.7	Illustrates the decreasing benefit of idealized personalized pricing over feature-based pricing as the number of incorporated model features increases. The numbers above the curve denote the scaled MAD of the unexplained noise, $\frac{E[\epsilon]}{2\mu}$, for every other k . We plot Logistic (left), Gumbel (middle), and Exponential (right) noise, respectively.	74
2.8	Illustrates the decreasing benefit of idealized personalized pricing over optimally segmenting into k groups for particular distributions (dotted lines), and our distribution-agnostic bound (solid line). Note the log-scales. . .	77
3.1	Depicted is a loot box offered by the popular video game <i>Dota 2</i> , which is called "treasure" in the game. The customer may purchase the loot box for \$2.49, after which they will receive one of the five items depicted on the screen.	92
3.2	The left panel shows an implementation of a loot box in the popular online game <i>Dota 2</i> . The red square highlights that it is a unique box as the loot box always allocates a unique item. The right panel shows an implementation of a loot box in the online game <i>PlayerUnknown's Battlegrounds</i> . It allows duplicates and is a traditional box.	100

3.3	The left panel shows an implementation of grand bundle selling in the online game <i>Brawlhalla</i> . All items may be unlocked for a one time payment of \$19.99 via the <i>All Legends Pack</i> (although customers may also buy “Mammoth coins” and subsets of the items). The right panel shows an implementation of separate selling in the online game <i>Arena of Valor</i> . In this game each item (character) can be individually unlocked for a single payment.	101
3.4	In the game <i>PlayerUnknown’s Battlegrounds</i> , the traditional box contains four classes of items: Mythic, Legendary, Epic, and Rare. The allocation probability for items of different classes varies, however items within the same class have the same probability.	111
3.5	The left panel shows an implementation of a loot box in the mobile game <i>Rise of the King</i> . Each box outputs 4 items. The right panel shows an implementation of a loot box in online game <i>League of Legends</i> . Customers are guaranteed to get one item, with the chance to get at most 2 bonus items within one box.	115
3.6	In the game <i>Dota 2</i> , players can trade in 6 unwanted items for a new loot box, plus 2000 shards, a form of in-game currency.	116
3.7	Normalized Prices for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.	120
3.8	Normalized Revenue for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.	121
3.9	Normalized Surplus for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.	121
A.1	134
A.2	Salop’s circle for $N=3$ products, divided into six equal regions.	135

A.3	These figures demonstrate the partition of the valuation space for a risk-neutral customer when $N = 2$. In both figures, $p = 4$. The left figure corresponds to a small discount for the opaque product and the right figure corresponds to a large discount for the opaque product. The four letters denote different buying behaviors of the customer under a (p, p^N) 1OPQ strategy and a SP strategy with price p	148
B.1	The left panel shows the bound in Theorem 2.5.2 versus the variance, using with $N = 200$ discretization points. The right panel plots the percent error in Theorem 2.5.2 when $M = .9, S = 10$ and $\frac{\sigma}{\mu} = 4$, as a function of N , the number of discretization points.	158
B.2	The left panel shows Theorem 2.5.2 with $N = 250$ points versus the scaled geometric mean. the right panel plots the percent error in Theorem 2.5.2 when $M = .9, S = 10$ and $\frac{G[V]}{E[V]} = .4$, as a function of N , the number of discretization points.	159

List of Tables

3.1	Possible Relations Between (UB) , (TB) , (GB) and (SS)	105
-----	--	-----

Acknowledgements

When I arrived to Columbia five years ago I wasn't sure I would make it through my first year, let alone the whole degree. What I hadn't accounted for was the kindness, generosity, and support from my friends, family, and the many brilliant individuals I have met here, without whom this disseration would not be possible.

First and foremost, I would like to thank my advisor, Adam. Adam introduced me to the problems that are central to this thesis, and guided me in this research since the very beginning. He taught me how to write papers. He helped me hone my (painfully bad) first talk, and my (marginally less painful) job talk, and every one of my talks in between. I know I was by no means the easiest Ph.D. student to advise, I really lucked out having you as an advisor. Thank you Adam for your support and patience, everything I know about being an academic I learned twice from you.

I would also like to thank my co-author and de facto mentor Vishal. Thank you for your patience, generosity, and for always upholding a standard of excellence in research. Whenever I was lost or stuck in my Ph.D., a sure way forward for me was to try and emulate you. I would like to thank my other coauthors Xiao and Ningyuan, who's talent and drive are an inspiration to me.

I would like to thank the other members of my committee. Thank you Jay and Vineet for your guidance and support, especially in this last year as I was on the job market. Thank you to Omar, your research is an inspiration to me. I would also to thank the many excellent professors I had at Columbia, especially Professor Stein, Professor Bienstock, and Professor Goldfarb. Your energy, enthusiasm, and love of

clever optimization have left a deep impression on me, one I hope to someday impart to my own students.

An enormous thank you to the front office, who have been unduly kind to me all five years I have been here. Without you all I literally would not have graduated, so thank you. More importantly, your kindness and warmth are what makes the office such a special place. In particular I would like to give special thanks to Liz, Kristen, Carmen, Gerald, Yoshi, and Jaya for being my friends. I looked forward to seeing you everyday and I will miss each of you when I am in Pittsburgh.

A big thank you to the many incredible Ph.D. students I have had the pleasure of sharing a department with over the years. I feel a little daunted at the prospect of trying to list them all but (in no particular order): Thank you Brian and Kevin, I wouldn't have gotten through the first year without you, I hope we can share another set of boots before I leave. Thank you to the members of the class above mine, especially Antoine, Chun, Francios, Jalaj, Suraj, Xinshang and Itai, I learned from and was inspired by you all. Special thanks to Irene, from whose STOC template I created my job talk. Thank you to my friends, Allen, Ryan, Wenbo, Xu, Wei, Raghav, Kumar, Jack, Yeqing, Apurv, Min, Randy, Wei, Nouri, Mali and everyone else I had the honor of sharing a department with. Working/goofing around with you all made the Ph.D. journey a pleasure.

I would like thank my little brother and my parents for always supporting me. It's to them I dedicate this thesis. Lastly I would like to thank my wonderful partner. Thank you Yi, you're the best discovery I've made here - sorry for all the late, late nights. Thank you all from the bottom of my heart.

To my parents, for everything.

Introduction

With the deluge of big data, retailers have more information about their customers and their preferences than ever before. Flush with rich, fine-grained market information, a host of new selling practices have proliferated in e-commerce retailing. In practice these strategies can take many novel forms, often involving many parameters or exploiting information asymmetry between the seller and the customers. In this dissertation we conduct a focused study of some of these emerging pricing schema and analyze their performance over traditional pricing technologies. In particular, we will take a closer look at three novel selling paradigms appearing in practice and in each case, place the strategies on firm theoretical grounding from which we can compare and contrast them with well studied pricing paradigms.

At a high-level the approach taken in this paper can be summarized as follows: first we observe a novel or alternative pricing scheme occurring in practice, for example opaque selling or personalized pricing. Next, we consider what traditional pricing structures the new pricing strategy is subverting, for example a single posted price or discriminatory pricing strategy. Finally, we place both the emerging pricing scheme and the traditional scheme in a grounded, random-utility maximizing framework and compare the expected revenue of each strategy against one another. In this thesis we apply this three step approach to opaque selling, personalized pricing, and loot box selling.

More generally, this thesis contributes to the larger field of revenue management and pricing, which uses quantitative analysis to determine what sort of products to

sell, how to sell those products, and at what prices to offer them. Our work concerns the latter two of these three, setting the right prices in the right way such that it maximizes a sellers revenue. While there are many possible factors to be taken into account when designing and analyzing revenue management systems, for instance the management of inventory or the effect of competition, and objectives for the system, for example social welfare or market penetration, this thesis focus solely on how to set prices for the goods so as to maximize revenue. This approach allows us to isolate the core differences between these emerging e-commerce technologies and their traditional counterpoints without muddying the story with overly complicated models. In relation to the deep literature on pricing, in this thesis we compare novel pricing paradigms against well studied strategies. In the process, we derive many new insights about well studied single price strategies and discriminatory pricing strategies.

In the remainder of this section we provide a short introduction to the framework employed in this work. We then flesh out the simplest pricing scenario and contrast it with the models studied in the next three chapters. Finally we summarize the contributions made in this thesis.

0.1 Introduction to Pricing for Revenue

Maximization

In this thesis we imagine markets of potential customers as being described by a distribution over valuations for the product(s). Each customers individual valuation(s) (i.e. willingness to pay) for the good(s) is represented as a draw from this distribution. Throughout the thesis we will use V to represent the customers valuations for the good(s), and F to be the distribution of V . Further, in this work we assume V is positive valued and admits a density f .

Now we describe how customers make purchasing decisions. A customer arrives to the seller and observes prices for the good(s). Based on the price and the customers drawn valuation(s), they choose to purchase the good that maximizes their *utility*, the difference between their valuation for the good and the price. When a customers valuation for all the goods are strictly less than their prices, we assume the customer will opt not to purchase anything. In the instance of ties (i.e. a customer's utility from two items is the same) we will make the standard assumption that a customer purchases the more expensive item. Often this assumption is without loss of generality. We call a customer behaving in this way *quasi-linear utility maximizing* and the overall model is known as *random utility maximizing* (RUM). We emphasize that, although the assumption of perfect rationality of a customer is quite strong, this model is ubiquitous across operations, business analytics and economics. All customers studied in this work will behave in this fashion.

Since markets are described by a distribution over valuations, F , and a customers purchasing decision depends solely on their drawn valuations and the fixed price; for any price p and market F there is an associated expected revenue where the expectation is taken over draws from F . By measuring this expected revenue we can assess the quality of a price. We will illustrate this framework in the next section, in the simple case of a single good.

0.2 Pricing a Single Good

To introduce the general paradigm considered in this thesis we will give a thorough treatment to the case of a retailer offering a single good. Suppose that the market is described by a positive valued random variable V drawn according to F . Then a customer offered the good at price p will purchase with probability $\Pr(V \geq p) = \bar{F}(p)$ where $\bar{F}(\cdot) := 1 - F(\cdot)$ is known as the survival function or the complementary

cumulative distribution function (cCDF). Since F describes the market, the expected revenue a seller will earn by pricing the good at p is

$$p\overline{F}(p) = \mathbb{E}[\mathcal{I}(V \geq p)p] := \mathcal{R}_{SP}(p),$$

where \mathcal{R}_{SP} is shorthand for the expected **revenue** of a **single price** p . Naturally the seller aims to maximize this revenue. When there is a single good, this maximization corresponds to $\max_p \mathcal{R}_{SP}(p)$. We refer to the maximum achievable revenue in this thesis as simply \mathcal{R}_{SP} . Now it is straight forward to characterize this optimal price in terms of the distribution. By checking first order conditions,

$$\frac{d}{dp} \mathcal{R}_{SP}(p) = \overline{F}(p) - pf(p) = 0$$

which implies the optimal price satisfies $p = \frac{\overline{F}(p)}{f(p)}$. The ratio $\frac{\overline{F}(p)}{f(p)}$ is the reciprocal of the well known as the hazard rate from reliability theory. If this ratio is decreasing in p (what is known as the monotone hazard rate condition (MHR)) it follows that the optimal price is unique. More generally, an optimal price can be found via standard one dimensional optimization techniques.

0.3 Pricing Beyond A Fixed Market and Single Good

In the three main chapters of this thesis we will be interested in the characterizing the revenue garnered by the optimal prices for a market. To provide some context for these results let's highlight a few nice features about the single good case that yielded a complete characterization of the achievable revenue, and contrast them with the models studied in the later three chapters.

1. (*Well Characterized Prices*) In the single item case the optimal price can be written down in implicit closed form, as a function of the distribution F . As we will see in Chapter 1 of this thesis, such nice implicit forms vanish when one attempts to generalize this setting beyond a single item i.e. to the multiple product case. When there are many competing goods, even when the underlying valuation distribution satisfies all our favorite assumptions (i.e. finitely supported, independent across items), just finding the optimal prices is intractably hard. In these cases, reasoning about the revenue of an optimal pricing strategies requires a softer touch, relying on more obscure properties the optimal prices must satisfy instead of an implicit characterization of the prices themselves.
2. (*Explicit Distributions*) The optimal price was easy to characterize via reference to the valuation distribution F . However in practice the *true* distribution of customer valuations is never known, instead only information about the distribution is known (for example the distributions mean and variance, but not its higher moments). This information induces a corresponding class of possible distributions matching the given information, we will refer to this class of possible distributions as Ω . In Chapter 2 we consider the subtler question of characterizing the optimal revenue when both the price and the distribution may vary, i.e.

$$\min_{F \in \Omega} \max_p p \overline{F}(p)$$

. Note the outer minimization is over a class of distributions Ω , which implies that whatever the solution is, it corresponds to a worst case guarantee on the revenue over all distributions matching the input data. Further, in this setting the number of parameters jumps from merely p to both p and a functional dependence on F .

3. (*Simple Customer Interaction*) Implicit in the characterization of the optimal

revenue in the previous section was the understanding that customers made one-time, instantaneous, irreversible decisions. From this assumption, we could calculate the expected revenue via a single expectation taken over F . In practice, the customer often has a repeated interaction with the seller, for example to collect multiple versions of the good or to return goods. In Chapter 3 we consider more expressive pricing models that allow for repeated interaction between the customer and seller. In this case the expected revenue is taken with respect to a stochastic process describing the customers interaction with the seller.

0.4 Summary of Contributions

We now summarize the main contributions of Chapters 1, 2 and 3. By collecting these results together, we can better contrast and compare them. Each of the following chapters will focus on a single emerging e-commerce technology and will be completely self contained.

In the first chapter we study the power of selling opaque products, i.e., products where a feature is hidden from the customer until after purchase. We consider models where traditional items are sold at a single price alongside opaque products corresponding to subsets of items and bench mark our opaque selling strategies against two common selling strategies: one which charges different prices for the items (discriminatory pricing), and one which charges a single price (single pricing). When customers are unit-demand and draw valuations exchangeably, we give a sharp characterization for when opaque selling outperforms discriminatory pricing, and characterize the maximal revenue lift opaque products can provide. This chapter is based on joint work with *Adam N. Elmachtoub*.

In the second chapter we turn our attention to personalized pricing strategies.

Access to high-quality customer information has fueled interest in personalized pricing strategies, i.e., strategies that predict an individual customer's valuation for a product and then offer them a customized price. While the appeal of personalized pricing is clear, it may also incur costs in the form of market research, investment in information technology, and branding risks. In light of these trade-offs, we study the value of personalized pricing over simpler pricing strategies, and provide various closed-form upper and lower bounds on the ratio that depend on simple statistics of the valuation distribution. This chapter is based on joint work with *Adam N. Elmachtoub* and *Vishal Gupta*.

Finally, in the last chapter we study the pricing and design of loot boxes in online games. In online gaming markets an increasingly popular way to sell in game items are via *loot boxes*, which are (random) bundles of virtual items, the contents of which are revealed after purchase. We consider how to design and price loot boxes, and compare their revenue with bundle selling and separate selling. We show that in an asymptotic regime, carefully designed loot box strategies can garner as much revenue as grand bundle selling while inheriting many nice properties of separate selling. We then extend these results to the case of multiple classes of items, in this setting we characterize the optimal allocation probabilities and prices, and salvage costs, where we show dominance over separate selling. In both cases we obtain closed form expressions for the asymptotic revenues. Finally, we numerically validate our results for moderately sized catalogs of items. This chapter is based on joint work with *Xiao Lei*, *Adam N. Elmachtoub*, and *Ningyuan Chen*.

Chapter 1

The Power of Opaque Products in Pricing

In this chapter we study the power of selling opaque products, the results of this section are joint work with *Adam N. Elmachtoub*. Opaque products are products where a feature (such as color) is hidden from the customer until after purchase. Opaque products, which are sold with a price discount, have emerged as a powerful vehicle to increase revenue for many online retailers and service providers that offer horizontally differentiated items. In the opaque selling models we consider, each of the items (colors) are sold at a single common price alongside opaque products which may correspond to various subsets of the items. We consider two types of customers, risk-neutral ones who assume they will receive a truly random item from the opaque product, and pessimistic ones who assume they will receive their least favorite item from the opaque product. We benchmark opaque selling against two common selling strategies: discriminatory pricing, where one explicitly charges different prices for each item, and single pricing, where a single price is charged for all the items.

We give a sharp characterization of when opaque selling outperforms discriminatory pricing. Namely, this result holds for situations where all customers are pessimistic, or the item valuations are supported on two points. In the latter case, we also show that opaque selling with just one opaque product guarantees at least 71.9% of the revenue from discriminatory pricing. We then provide upper bounds on the potential revenue increase from opaque selling strategies over single pricing, and describe cases where the increase can be significantly more than that of discriminatory pricing. Finally, we provide pricing algorithms alongside a numerical study to

assess the power of opaque selling under various valuation distributions.

1.1 Introduction

An opaque product is a product where one or more features (such as color, brand, or location) are hidden from the customer until after the purchase is made. In recent years, several online retailers have begun selling opaque products. For example, Amazon.com offers various colors of Swingline staplers alongside a “colors may vary” option, which is an opaque product over the various colors (see Fig. 1.1). In another example, SwimOutlet.com offers various styles of Nike swimsuits, as well as a “Grab Bag” over all the different styles offered (see Fig. 1.2).

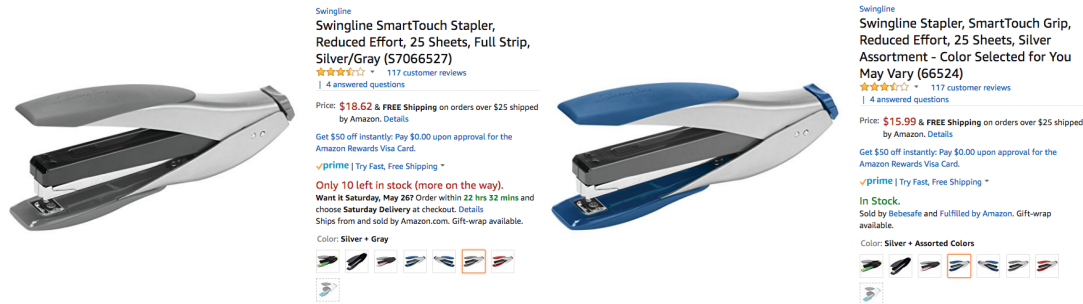


Figure 1.1: Swingline offers their “SmartTouch” staplers on Amazon.com traditionally alongside a “colors may vary” option i.e., a single opaque product where the color of the stapler is hidden until after purchase. The opaque product (right) is offered at the discounted price of \$15.99, which is \$1.63 less than the traditional price (left) of \$18.62.

In both of these examples, customers who purchase the opaque product sacrifice exact knowledge of the item they will receive in exchange for a price discount. This allows the seller to price discriminate between customers with strong and weak preferences, and therefore earn more revenue. The goal of this work is to showcase the power of opaque selling compared to more traditional price discrimination tactics, and quantify the potential extra revenue that a seller can obtain.

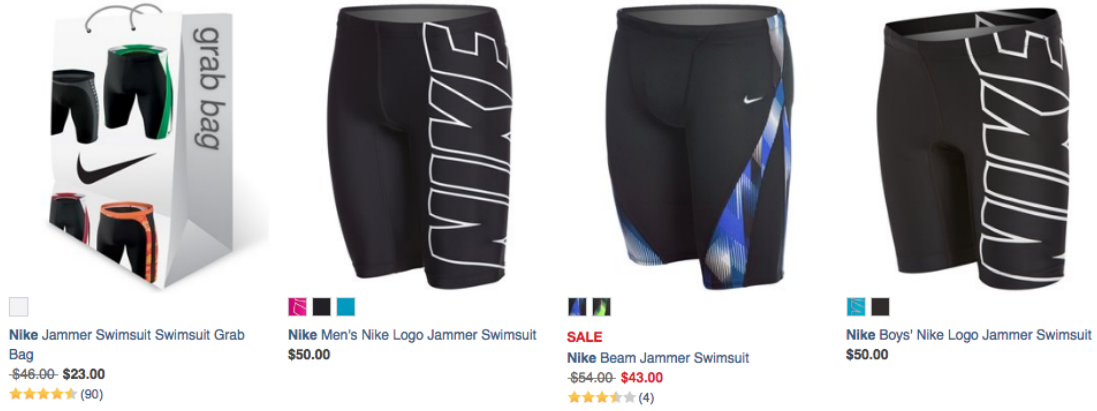


Figure 1.2: SwimOutlet.com offers various styles of Nike brand swim trunks for prices between \$43.00-\$50.00 alongside an opaque “Grab Bag” for \$23.00.

In our framework, we consider a seller that has N items available for sale, each of which are similar but may differ in a secondary attribute such as color or style. Customers draw valuations for each of the items from a joint probability distribution that is known to the seller. We focus on the class of exchangeable distributions which naturally represent horizontally differentiated items, as the marginal distribution of each item is identical. Special cases include i.i.d. valuations, Hotelling model, and Salop’s circle (when the number of items is fewer than 3). In the absence of opaque products, customers simply choose the item which maximizes their utility, i.e., their valuation for the item minus its price. No item is purchased in the case where the utility from all items is negative.

Interestingly, when the valuation distribution is exchangeable (even i.i.d.), the optimal pricing strategy in this model may use different prices for different items (Chawla et al. [30]). This strategy, that we refer to as discriminatory pricing, is a natural benchmark for our opaque selling strategies. Due to symmetry, discriminatory pricing arbitrarily chooses some items to have high prices. This may be particularly problematic when certain items (colors) are correlated with demographic information such as race or gender. In some settings, the items are often constrained to have the same price by the manufacturer or by the seller to ensure impartiality to customers.

Thus another natural benchmark to consider is the best single price strategy.

We now carefully describe our opaque selling strategy, where the seller offers opaque products in addition to offering the N items. Specifically, an *opaque product* is an explicit subset of items from which a customer will receive one item upon purchase. An opaque selling strategy can offer all possible $2^N - N - 1$ opaque products in addition to the N original items. For practicality and tractability of the model, we assume that opaque products corresponding to subsets of the same size must have the same price. Moreover, we impose a restriction that all prices of the items must be the same. This exact structure is employed by Eurowings (see Fig. 1.3 and Post and Spann [90]) which sells round trip tickets to opaque destinations. In this setting, customers may narrow down the possible destinations in exchange for an increased price. As a result, an opaque selling strategy is parameterized by N prices, similar to a discriminatory pricing strategy. However, customers interested in an item or opaque products of the same size always pay the same price, which prevents the opaque strategy from arbitrarily discriminating against a particular type of customer. *In essence, an opaque selling strategy balances the impartiality of a single price strategy with the price discrimination capability of discriminatory item pricing.*

In order to study opaque selling strategies, we must also specify how customers value an opaque product. We consider two approaches motivated by realistic interpretations of opaque products, which we call pessimistic and risk-neutral customers. In practice, the seller never reveals the probabilities of receiving individual items in an opaque product, leaving the customers to formulate their valuations based on their judgment. A customer is said to be *pessimistic* if they value an opaque product as the *minimum* of their valuations among the corresponding subset of items. A pessimistic customer is risk-averse and essentially wants an ex-post guarantee that purchasing the opaque product maximizes their utility. Such an assumption is natural when some of the items are infeasible for the customer (mismatched colors or incompatible des-

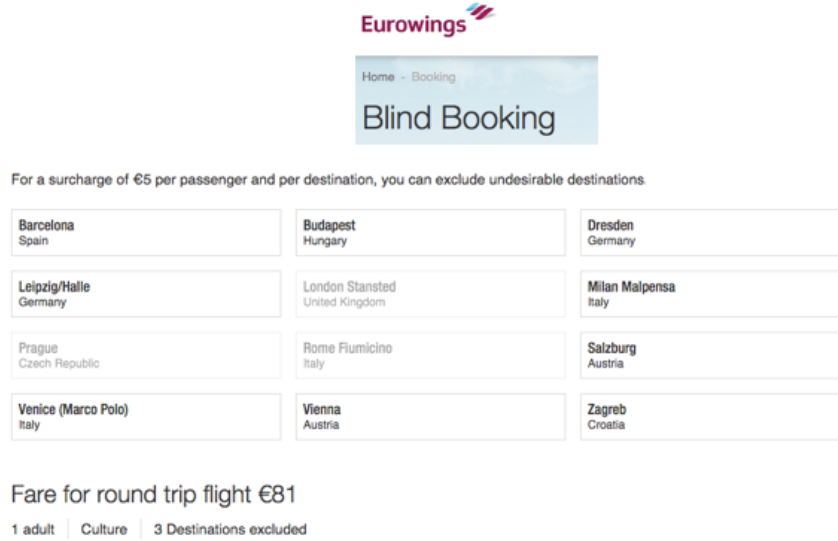


Figure 1.3: Eurowings.com uses an opaque selling strategy to offer airline tickets with a base price of €66.00. Here the destination of the flight is opaque, with $N = 12$ possible destinations. The site allows the customer to exclude as many destinations as they desire, each for a price of €5.00. In the figure above, three destinations are excluded and the price for the desired opaque product is €81.

tionation). Another reason a customer may be pessimistic is that he is fundamentally mistrustful of the seller's motives, and believes the seller will allocate the product he desires least.

A customer is said to be *risk-neutral* if they value an opaque product as the *average* of their valuations among the corresponding subset of items. A risk-neutral customer is essentially optimistic, and believes that the seller is impartial in the sense that the probability of receiving any item in an opaque product is uniformly distributed. Although not always true, it has been the dominant assumption in the literature and may model customer behavior well in some applications. We study the power of opaque selling in markets that simultaneously allow a mixture of both customer types, where α denotes the probability a customer is pessimistic.

We next outline our contributions, which formally describe conditions under which opaque selling performs well with respect to discriminatory and single pricing strategies.

1. We give a sharp characterization of when opaque selling dominates the optimal discriminatory pricing strategy. In particular, opaque selling is *guaranteed* to provide more revenue than discriminatory pricing when the valuations are drawn from an exchangeable distribution and either of the following conditions hold: (i) the market is homogeneously pessimistic or (ii) the valuations can only take two values (high or low). When neither condition holds, either strategy may be superior to the other. One surprising consequence is that the seller may actually benefit when customers are pessimistic and hence value opaque products in a worst-case manner.
2. In the important special case when valuations are drawn from an exchangeable distribution and can only take two values, we show a single opaque product can always guarantee *at least 71.9%* of the revenue from discriminatory pricing. This result is independent of what fraction of the customers are pessimistic and risk-neutral.
3. We then show that in α -mixed markets, opaque selling can earn *up to and at most a factor of αN* more than the best single pricing. When this bound is tight, this revenue increase is exponentially more than the revenue increase provided by any discriminatory pricing. We compliment this result by showing that in the restricted case of i.i.d. valuations, this gap falls to constant factors. We also show that offering just a single opaque product increases revenue by *up to and at most a factor of $2 - (1 - \alpha)\frac{N-1}{N}$* over the best single pricing.
4. We perform a numerical study which bears out our results for several typical distributions. To conduct the study we derive an *efficient* algorithm for finding the optimal prices when N is small. We empirically observe that the seller earns more revenue as α , the fraction of pessimistic customers, increases. We also see up to a 5% increase in revenue using opaque selling compared to discriminatory pricing. Finally, when N is small, we also observe that a single opaque product

earns nearly as much revenue as a general opaque strategy.

In summary, our results provide strong evidence that opaque selling is a customer-friendly alternative to discriminatory pricing, often with comparable or higher revenues. Next, we compare our results to previous work and connect our ideas to related literature.

Literature Review

Our work connects into several streams of literature across operations, marketing, economics, and computer science. We first review literature on monopolistic sellers offering opaque products. For example, the parallel works of Jiang [75] and Fay and Xie [53] both consider opaque selling frameworks when customers have valuations drawn from a Hotelling or Salop’s circle choice model. Note that this assumes perfect correlation between the items, and may not necessarily represent customer behavior well although it does fall into the exchangeable distribution assumption. Both works provide conditions for when opaque selling can have strictly positive increase in profit over single pricing strategies. Fay and Xie [53] also show that opaque products can be used to hedge against possibly incorrect demand estimation. Our work does not make any assumption about the valuation distribution, and benchmarks against discriminatory pricing.

A separate stream of work has shown the power opaque products for managing capacity and inventory. Gallego and Phillips [59] and Gallego et al. [60] considers the notion of a flexible product in revenue management, where customers who buy the flexible option are allocated a product after the completion of the time horizon. Fay and Xie [54] show how to use opaque products to protect inventory when one of the two items is strictly preferred over the other by all customers, and Xiao and Chen [103] provide dynamic programming algorithms to decide when to use opaque products. Elmachoub and Wei [50] quantify the value of opaque products in real-

time inventory management environments, and provides a framework for computing choice probabilities when $N = 2$ and customers are pessimistic. In our work, we avoid any notion of cost and focus purely on the price discrimination effect offered by opaque products.

There are several works on opaque products when used among competitors (Shapiro and Shi [96], Jerath et al. [74]), in name-your-own-price channels (Chen et al. [34], Huang et al. [72]), in empirical analysis (Xie et al. [104], Granados et al. [63]), and in queueing systems (Xu et al. [105], Geng [62]). Post and Spann [90] and Post [89] consider settings where multiple opaque products are offered simultaneously. We also mention a stream of work in economics that considers optimal mechanism design with opaque products, along the lines of Pavlov [86], Balestrieri et al. [12], and Balestrieri and Izmalkov [11].

Our results related to purely risk-neutral markets connects to a stream of literature on pricing with lotteries. A lottery, as described in the literature, is a probability distribution over the items that is sold by the seller and announced to the customer. Customers are risk-neutral and use the expected valuation of the lottery when deciding what to buy. If all customers are risk-neutral, then our opaque selling strategy can be thought of as a special case of lottery pricing where the items are allocated uniformly at random. Under arbitrary valuation distributions, Briest et al. [24] and Hart and Nisan [67] show that lottery pricing can earn infinitely more revenue than any discriminatory pricing when $N \geq 3$ and $N = 2$, respectively. When customers draw their valuations independently, Chawla et al. [32] show the optimal lottery pricing are at most four times discriminatory pricing.

More generally, our work is related to a growing literature in computer science on algorithmic pricing. When valuations are independent, Chawla et al. [30] and Chawla et al. [31] give constant factor approximations to the optimal discriminatory pricing by exploiting a surprising connection to a 1 item, N bidder auction. Cai and Daskalakis

[26] provide an additive polynomial-time approximation scheme when valuations are bounded and drawn independently, while Chen et al. [37] show the problem is NP-Hard even when valuations are drawn i.i.d. or drawn independently and supported on three points. The class of exchangeable valuations was also considered in Daskalakis and Weinberg [46] in the context of optimal auction design.

We note that our work resembles that of bundle pricing on the surface due to the nature in which items are aggregated into opaque products, although bundling results generally assume customers are interested in purchasing multiple items. The one exception is that of Briest and Roglin [22] who frame opaque products as ‘unit demand bundles’ and provide hardness results. Finally, our work fits in parallel to recent work on simple mechanisms for difficult multi-dimensional problems in auctions (Celis et al. [28]) and bundling (Ma and Simchi-Levi [79], Abdallah et al. [3]).

1.2 Selling Models

We now formally describe the selling models that we study throughout the work. We consider a seller who has $N \geq 2$ items for sale, described by the set $\mathcal{N} := \{1, 2, \dots, N\}$. The seller may also offer one or more opaque products, each of which is described by a subset $S \in 2^{\mathcal{N}}$ where $|S| \geq 2$. The seller simultaneously offers the items and potentially some number of opaque products to a utility-maximizing, unit-demand customer. (Note that this is equivalent to selling to many customers with no inventory constraints.) The customer has a nonnegative random valuation for each item i denoted by V_i , and the joint valuation $V = (V_1, V_2, \dots, V_N)$ is drawn from a known joint distribution F . For every selling model we consider, the customer maximizes his own utility, which is the valuation of the item or opaque product purchased minus its price. If no item or opaque product results in a nonnegative

utility, then the customer does not purchase anything. In the case where the customer has multiple options that provide maximum utility, we make the standard assumption, without loss of generality, that the customer purchases the product with the highest price (see [37] for detailed discussion of tie-breaking rules in this context). When there are multiple products with the same price providing maximum utility, we assume the customer breaks ties arbitrarily.

We note that the notion of valuation and utility described thus far does not extend in an obvious way to opaque products. That is, the way a customer values an opaque product depends on the customer’s belief about the seller’s allocation mechanism and the customer’s risk preferences. Next, we describe two natural frameworks for capturing valuations of opaque products, and each customer shall behave according to one of these two frameworks.

Valuations for Opaque Products

We model the customer’s valuation for an opaque product as a function over the valuations of the items the opaque product can return. We consider two natural assumptions for how to model customer behavior with respect to opaque products, which we call *pessimistic* and *risk-neutral*. For any subset of items $S \in 2^N$, we let V^S denote the random valuation of the opaque product corresponding to S . For pessimistic customers, V^S is the minimum over all the valuations in S , i.e.,

$$V^S = \min_{i \in S} \{V_i\}$$

and for risk-neutral customers, V^S is the average over all the valuations in S , i.e.,

$$V^S = \frac{\sum_{i \in S} V_i}{|S|}.$$

We assume that a customer is pessimistic with probability α and risk-neutral with probability $1 - \alpha$. We let X_α denote the random variable corresponding to the customer type.

In practice, the seller *never* announces the allocation probabilities for an opaque product, forcing the customer to form his own beliefs. A pessimistic customer believes the seller will allocate the product that is desired least by the customer. Given that the allocation probabilities are entirely unknown, this corresponds to a customer placing a worst-case allocation distribution on the outcome of the opaque product. The pessimistic preference also captures another important and practical situation where even if the customers know the opaque allocation probabilities, they are extremely risk-averse. In other words, the customer wants their purchasing decision to be ex-post optimal, i.e., there is no regret even after the item in the opaque product is revealed. This particular situation can arise when customers know that certain items provide no value, which can happen when particular colors or flight destinations are completely undesirable (see Figures 1.2 and 1.3).

A risk-neutral customer believes that the seller will allocate the items in the opaque product uniformly at random, which is an optimistic belief. With respect to this fair allocation, he is also risk-neutral in his valuation of the opaque product. Thus, a risk-neutral customer simply averages their valuations across the product, even though the allocation is most likely not uniformly at random. With limited information, it is natural for some customers to form this valuation, in particular when the valuations of each item are reasonably close together (in which case the difference between risk-neutral and pessimistic is small). We also note that the risk-neutral assumption has been the primary focus in the literature (Gallego and Phillips [59], Fay and Xie [53], Jerath et al. [74]), while the pessimistic case has not been studied to our knowledge.

Finally, we highlight that the delineation between pessimistic and risk-neutral

customers is quite important from a geometric and technical perspective. Fig. 1.4 illustrates this distinction in the case where $N = 2$. Note the shape and size of the valuations regions where customers purchase the opaque product are quite different, which explains the dependence on α in our analysis.

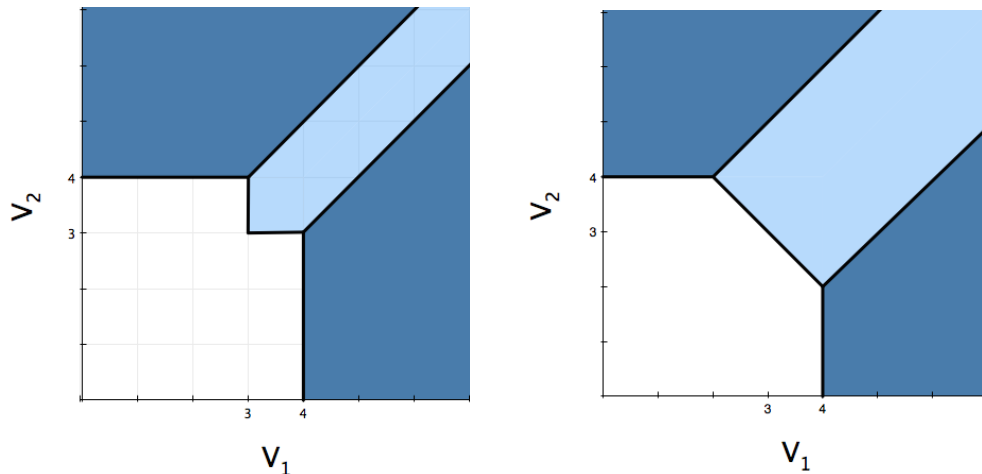


Figure 1.4: Above are two valuation spaces for opaque selling strategies with prices $(p, p^2) = (4, 3)$. Left is the valuation space for pessimistic customers, right is valuation space for risk-neutral customers. The darkened regions correspond to customer valuations that yield purchases of an item at a price of 4. The lighter regions correspond to purchases of the opaque product which has a price of 3. The unshaded regions correspond to valuations that yield no purchase.

Selling Strategies

We now describe four specific selling strategies that we use throughout the work. For notational convenience and improved exposition of this subsection, we assume that F is continuous to avoid tie-breaking scenarios (which would go to the highest price option w.l.o.g.). In the *single price selling model* (SP), the seller offers all N items all at the same price. In other words, the price of item i is the same for all $i \in \mathcal{N}$. We refer to this single price as p . We denote $\mathcal{R}_{SP}^F(p)$ and \mathcal{R}_{SP}^F as the expected revenue using single pricing with joint distribution F under price p and the optimal price,

respectively. More formally,

$$\mathcal{R}_{SP}^F(p) = p \Pr(\max_{i \in \mathcal{N}} \{V_i\} \geq p) \text{ and } \mathcal{R}_{SP}^F = \max_p \mathcal{R}_{SP}^F(p).$$

In the *discriminatory pricing model* (DP), the prices may differ between the items. Without loss of generality, we always relabel the indices so that $p_1 \geq p_2 \geq \dots \geq p_N$. We denote the vector of prices as \vec{p} . We note that even when valuations are i.i.d., discriminatory item pricing may provide strictly more revenue than single pricing strategies (Chawla et al. [30]). However, discriminatory pricing may be difficult for customers to accept (especially when valuations are i.i.d.), and for similar reasons may be infeasible for the seller due to business constraints. We denote $\mathcal{R}_{DP}^F(\vec{p})$ and \mathcal{R}_{DP}^F as the expected revenue under F using prices \vec{p} and the optimal item pricing, respectively. More formally,

$$\mathcal{R}_{DP}^F(\vec{p}) = \sum_{i \in \mathcal{N}} p_i \mathbb{P} \left(V_i - p_i \geq \max_{j \neq i} \{V_j - p_j, 0\} \right) \text{ and } \mathcal{R}_{DP}^F = \max_{\vec{p}} \mathcal{R}_{DP}^F(\vec{p}).$$

Although DP seems at first unnatural and counterintuitive when valuations are i.i.d., the revenue function creates a natural tension to segment the market and capture high valuation customers without sacrificing market size. Every time an item is priced high, selling another item at a low price becomes more valuable since its market share will increase. Although DP can provide more revenue in many i.i.d. settings, including two point (high-low) distributions, it may not be beneficial in other settings. For example, when i.i.d. valuations correspond to a multinomial logit (MNL) choice model, the ‘constant markup property’ (Anderson et al. [4]) implies SP is optimal. We provide a longer primer delving further into DP in Appendix A.1.

In the *single opaque selling model* (1OPQ), the seller offers only one opaque product associated with the set \mathcal{N} at a price p^N , alongside the traditional items all at a fixed price p . This model is important and used in practice due to its simplicity,

impartiality, and ease of implementation. We denote $\mathcal{R}_{1OPQ}^{F,\alpha}(p, p^N)$ and $\mathcal{R}_{1OPQ}^{F,\alpha}$ as the expected revenue under F , in α -mixed markets, using prices (p, p^N) and the optimal pricing, respectively. Formally,

$$\begin{aligned}
\mathcal{R}_{1OPQ}^{F,\alpha}(p, p^N) &= \alpha p \mathbb{P} \left(\max_i \{V_i - p\} \geq \max\{V^N - p^N, 0\} \mid \text{Pessimistic} \right) \\
&\quad + \alpha p^N \mathbb{P} \left(V^N - p^N > \max_i \{V_i - p, 0\} \cap V^N - p^N \geq 0 \mid \text{Pessimistic} \right) \\
&\quad + (1 - \alpha) p \mathbb{P} \left(\max_i \{V_i - p\} \geq \max\{V^N - p^N, 0\} \mid \text{Risk Neutral} \right) \\
&\quad + (1 - \alpha) p^N \mathbb{P} \left(V^N - p^N > \max_i \{V_i - p, 0\} \cap V^N - p^N \geq 0 \mid \text{Risk Neutral} \right) \\
&= E_{X_\alpha} \left[p \mathbb{P} \left(\max_i \{V_i - p\} \geq \max\{V^N - p^N, 0\} \mid X_\alpha \right) \right. \\
&\quad \left. + p^N \mathbb{P} \left(V^N - p^N > \max_i \{V_i - p, 0\} \cap V^N - p^N \geq 0 \mid X_\alpha \right) \right] \\
\text{and } \mathcal{R}_{1OPQ}^{F,\alpha} &= \max_{p, p^N} \mathcal{R}_{1OPQ}^{F,\alpha}(p, p^N)
\end{aligned}$$

The first equation is derived by considering four events that generate revenue. The first event is that a customer is pessimistic, which has probability α , and has a valuation for their favorite item, $\max_i \{V_i - p\}$, which is larger than the opaque product and 0, $\max\{V^N - p^N, 0\}$. In this event, the customer buys an item and the seller earns p . The remaining events are similar and enumerate the other cases where an opaque product is purchased and/or the customer is risk-neutral.

In the general *opaque selling model* (OPQ), the seller offers all possible opaque products, alongside the items which are offered at a single price p . For simplicity, tractability, and impartiality, opaque products of the same cardinality are assigned the same price (see Figure 1.3 for an example of this exact scenario). That is for all $S, S' \in 2^N$ s.t. $|S| = |S'| \geq 2$, the opaque products corresponding to the subsets S and S' must have the same price. For subsets of size k , the corresponding price is p^k , and the vector of the $N - 1$ opaque product prices is denoted by \vec{p} . We denote $\mathcal{R}_{OPQ}^{F,\alpha}(p, \vec{p})$ and $\mathcal{R}_{OPQ}^{F,\alpha}$ as the expected revenue under F , in α -mixed markets, using

prices (p, \vec{p}) and the optimal pricing, respectively. More formally,

$$\begin{aligned} \mathcal{R}_{OPQ}^{F,\alpha}(p, \vec{p}) &= E_{X_\alpha} \left[p \mathbb{P} \left(\max_{i \in \mathcal{N}} \{V_i - p\} \geq \max_{S \in 2^\mathcal{N}, |S| \geq 2} \{V^S - p^{|S|}, 0\} \mid X_\alpha \right) \right. \\ &\quad \left. + \sum_{S \in 2^\mathcal{N}, |S| \geq 2} p^{|S|} \mathbb{P} \left(V^S - p^{|S|} \geq \max_{S' \in 2^\mathcal{N}, |S'| \geq 2} \{V^{S'} - p^{|S'|}, 0\} \cap V^S - p^{|S|} \geq \max_{i \in \mathcal{N}} \{V_i - p\} \mid X_\alpha \right) \right] \\ \text{and } \mathcal{R}_{OPQ}^F &= \max_{p, \vec{p}} \mathcal{R}_{OPQ}^F(p, \vec{p}). \end{aligned}$$

In this expression, the expectation is taken with respect to the customer type, which affects the opaque product valuations. The first summand corresponds to the revenue in the case where an item is bought, and the remaining summands corresponds to the revenues of the opaque products. Note an opaque product is sold only if it provides nonnegative utility and has more utility than all of the items and other opaque products.

We note that implementing the OPQ strategy in practice is simple, despite the exponentially large number of products (see Figure 1.3). The seller simply displays the price ladder corresponding to the size of the opaque product purchased, and then users simply select (click) their top k products if they chose to purchase an opaque product of size k . We also note that OPQ and 1OPQ are equivalent when $N = 2$. For readability we often omit the superscripts F and α when they can be inferred from context. In general, subscripts always refer to item prices (p_i) and superscripts refer to opaque product prices ($p^{|S|}$).

Valuation Distributions

In this work we focus on the broad class of *exchangeable* valuation distributions, which generalizes i.i.d. valuations to distributions that allow for structured symmetric correlation between items. The formal definition is presented below.

Definition 1.2.1. *We call the random valuation vector $V = (V_1, \dots, V_N)$ exchangeable if every permutation of the item valuations results in the same joint distribution. The corresponding distribution is also said to be exchangeable in this case.*

Exchangeable valuation distributions are a natural model for horizontally differentiated items as they allow for individual preferences between items, but enforce a distributional symmetry as the items are all alike. One important example is the Hotelling model which has been the primary focus of previous works (Fay and Xie [53], Jerath et al. [74]) which focus on two item scenarios. Salop’s circle is a generalization of the Hotelling model for more than two items, and is a standard choice model for capturing horizontal differentiation (Salop [93], Fay and Xie [53]). When $N = 3$, Salop’s circle model is an exchangeable distribution. For $N \geq 4$, a more general, but complex, notion of exchangeability is needed to capture Salop’s circle. We provide this definition in Appendix A.2 and note that many of our results extend to this more general definition. For ease of exposition, we shall focus on Definition 1.2.1 throughout the paper.

1.3 The Power of Opaque Products

In this section, we focus on the revenue from the general opaque (OPQ) strategy when item valuations are drawn from any exchangeable distribution. In Sections 1.3 and 1.3, we provide conditions for when the expected revenue of OPQ is guaranteed to exceed that of discriminatory pricing (DP). When neither of these conditions hold, there is no dominance in either direction, and we supply counterexamples (valuation distributions) where discriminatory pricing is better. In Section 1.3, we quantify how much more revenue OPQ selling strategies can potentially earn over single pricing (SP), and show that the extra revenue garnered by OPQ strategies can be on the order of αN more than DP. In the special case where item valuations are i.i.d., we

show this gap collapses to a constant factor.

Benchmarking against Discriminatory Pricing

We now characterize when OPQ is guaranteed to garner more revenue than DP. In particular, we show that when all customers are pessimistic or when valuations can take only two values (high or low), opaque selling is guaranteed to earn more revenue than discriminatory pricing. In Example 1.3.1, we give a valuation distribution where neither condition holds and $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$. This counterexample assumes valuations can take three values, and assumes that $0 \leq \alpha \leq .85$. Next, we formally state our result in Theorem 1.3.1 and defer the proof to Section 1.3.

Theorem 1.3.1 (When OPQ dominates DP). *Assume customers are α -mixed and draw their valuations from an exchangeable distribution. If (i) $\alpha = 1$ or (ii) the item valuations take only two values, then opaque selling dominates discriminatory pricing, i.e.,*

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}.$$

Interpretation and Implications of Theorem 1.3.1: While restricted, both cases of Theorem 1.3.1 where the dominance result holds represent situations of significant interest. When α is near 1, most customers assume a worst-case behavior with respect to opaque product allocation. This situation arises in markets where opaque products have been recently introduced and there is no information for customers to be had. For markets where customers tend to value a particular item (color or destination) as unacceptable, this pessimistic behavior may also be common. A trivial implication of Theorem 1.3.1(i) is that $\mathcal{R}_{OPQ} \geq \alpha \mathcal{R}_{DP}$, which follows from simply ignoring the revenue from all risk-neutral customers. Thus in highly pessimistic markets where α is close to 1, an opaque selling strategy is guaranteed to preserve almost all the gains from discriminatory pricing, and potentially earn even more.

Further, we show in Corollary A.2.1 that Theorem 1.3.1(i) extends to another important class of distributions for horizontally differentiated items known as Salop’s circle, often used as a standard tool in the literature. We provide a short primer on Salop’s circle model along with the proof of Corollary A.2.1 in Appendix A.2.

When valuations are supported on two points, the market is highly differentiated and has binary ‘high/low’ valuations for the items. This setting has been the subject of Fay and Xie [53], Huang and Yu [71] in the literature on opaque selling. Note that the case of binary valuations is exactly when discriminatory pricing is most profitable compared to single pricing: Dutting and Klimm [49] shows that for every N , there exists an i.i.d. two point distribution such that $\mathcal{R}_{DP} = (2 - \frac{1}{N}) \mathcal{R}_{SP}$ and that this is the largest possible revenue gap. Thus in markets where retailers would be most inclined to consider discriminatory pricing strategies, an opaque selling strategy is even more profitable.

It is important to note that when the conditions of Theorem 1.3.1 do not hold, that either DP or OPQ may be preferred depending on the market assumptions. Thus it is worth noting that OPQ may have other advantages over DP. For example, discriminatory selling can be unnatural and undesired by customers in particular in the settings we consider where the items only differ superficially (hence exchangeable distributions). Charging different prices for what are essentially equivalent products may increase revenue, however it may be perceived as unfair by customers (and cause strategic behavior) or even disallowed by manufacturers altogether. In contrast, the OPQ strategy is impartial and will never result in a customer paying more simply for liking a particular item (color). Collectively, we believe these arguments show that opaque selling should always be considered as an alternative to discriminatory pricing, and in many cases may result in more profit.

Sharpness of Theorem 1.3.1. Both the exchangeability and sufficient conditions ($\alpha = 1$ or two point valuations) for Theorem 1.3.1 are critical for the result to hold. In Example 1.3.1, we construct a simple three-point distribution from which item valuations are drawn i.i.d. and $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$ for any $\alpha \leq .85$. Thus Example 1.3.1 precludes generalizing Theorem 1.3.1 for situations beyond two point valuations and purely pessimistic markets. Surprisingly, it also implies that when $\alpha = 1$, the revenue from OPQ may be higher than when $\alpha = 0$. In other words, the seller may actually *benefit* from customers adopting a pessimistic attitude towards opaque products, as this helps segment the market more favorably. In Example 1.3.2, we describe a valuation distribution that is not exchangeable and results in $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$

Example 1.3.1 (When Assumptions (i) and (ii) Do Not Hold). *For $N = 2$, we construct a three point distribution F where, when customers are risk-neutral, the optimal discriminatory selling strategy earns strictly more revenue than an opaque strategy. Let $\alpha = 0$ and suppose i.i.d valuations for two items drawn according to,*

$$V_i = \begin{cases} 0 & : w.p \ 8/27 \\ .1 & : w.p \ 2/3 \\ .9 & : w.p \ 1/27 \end{cases}$$

Then using the algorithm described in Theorem 1.5.1 we can compute $\mathcal{R}_{SP} = \mathcal{R}_{OPQ} = 0.091220... < 0.0913$ achieved by pricing both items at 0.1. However $\mathcal{R}_{DP}(0.9, 0.1) = 0.1$, 11% more revenue than the optimal opaque selling strategy.

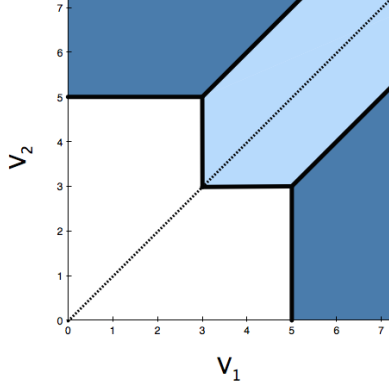
Further, it can be computed that when $\alpha > .69231$, the optimal opaque pricing switches from $(.1, .1)$ to a mixed pricing $(.1, .9)$ earning revenue $\approx \alpha(1.08779) + (1 - \alpha)(.0517146)$. The revenue from this optimal mixed opaque selling strategy overtakes the revenue from discriminatory pricing when $\alpha > .846$. Thus for $\alpha < .69231$, $\mathcal{R}_{SP} = \mathcal{R}_{OPQ} < \mathcal{R}_{DP}$. For $\alpha \in (.69231, .846)$, $\mathcal{R}_{SP} < \mathcal{R}_{OPQ} < \mathcal{R}_{DP}$. For $\alpha > .846$, $\mathcal{R}_{SP} < \mathcal{R}_{DP} < \mathcal{R}_{OPQ}$. \square

Example 1.3.2 (When Exchangeability Does Not Hold). *Consider a market where $N = 2$ and $\alpha = 1$, and where valuations for two items are drawn independently from V_1 , which is two times a Bernoulli r.v. with probability $1/2$, and V_2 which is distributed as a Bernoulli r.v. with probability $1/2$. Since V_1 and V_2 are independent but not identical, the market is therefore not exchangeable. However, by a simple enumeration, one can see that $\mathcal{R}_{DP}(2, 1) = \frac{5}{4}$ whereas $\mathcal{R}_{1OPQ} \leq 1$. \square*

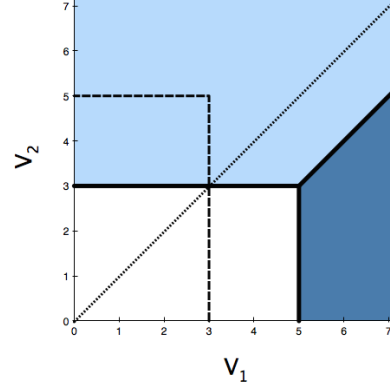
Geometric Proof of Theorem 1.3.1 when $N = 2$ and $\alpha = 1$. Before delving into the formal proof in Section 1.3, we provide some geometric intuition in the special case when $N = 2$ and $\alpha = 1$. Suppose the optimal discriminatory pricing uses prices (p_1, p_2) with $p_1 > p_2$. We show that an opaque selling strategy with prices $(p, p^2) = (p_1, p_2)$ exceeds the revenue of the optimal discriminatory pricing. Fig. 1.5a and Fig. 1.5b show the different purchase behaviors under OPQ and DP, respectively, where a darker color corresponds to a more expensive customer purchase. Due to exchangeability, it is then visually clear that the following are all equal: (i) the revenue of OPQ conditioned on $V_1 \geq V_2$, (ii) the revenue of OPQ conditioned on $V_2 \geq V_1$, and (iii) the revenue of DP conditioned on $V_1 \geq V_2$. To complete the proof, we claim that the revenue of DP conditioned on the event $V_1 \geq V_2$ is at least the revenue of DP conditioned on the event $V_2 \geq V_1$. If this were not the case, then reducing p_1 to p_2 would increase the revenue in the event that $V_1 \geq V_2$ without changing the revenue in the event $V_2 \geq V_1$, which would contradict the optimality of (p_1, p_2) .

One interesting consequence of this geometric argument is that, when $N=2$ and $\alpha = 1$, $\mathcal{R}_{DP} \leq \frac{\mathcal{R}_{OPQ} + \mathcal{R}_{SP}}{2}$. Suppose that $\mathcal{R}_{DP} = (1 + \gamma)\mathcal{R}_{SP}$, for some $\gamma > 0$. Then rearranging $\mathcal{R}_{DP} \leq \frac{\mathcal{R}_{OPQ} + \mathcal{R}_{SP}}{2}$ gives

$$\mathcal{R}_{OPQ} \geq \frac{1 + 2\gamma}{1 + \gamma} \mathcal{R}_{DP}$$



(a) Valuation space for pessimistic customers facing a OPQ strategy with $(p, p^2) = (5, 3)$. Note the purchasing behavior is symmetric across the line $V_2 = V_1$.



(b) Valuation space for a customer facing a DP strategy with $(p_1, p_2) = (5, 3)$. Note that below the line $V_2 = V_1$, the purchasing behavior in (a) and (b) are identical.

Figure 1.5: The valuation space and purchasing behaviors for a pessimistic customer facing OPQ and DP selling strategies respectively. Customers with valuations in the darkened regions buy at price 5 in both figures. Customers with valuations in the lightly shaded regions buy at price 3 (i.e., purchase the opaque product or item 2, respectively). Customers in the unshaded region do not purchase.

which implies the inequality in Theorem 1.3.1 is strict whenever $\mathcal{R}_{DP} > \mathcal{R}_{SP}$. In Corollary 1.3.1 we expound on this observation to show more generally, whenever the conditions of Theorem 1.3.1(i) are met, and $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, it follows that $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$.

When the conditions of Theorem 1.3.1(ii) hold, no such result can be shown. When α is small, there are cases when $\mathcal{R}_{DP} = \mathcal{R}_{OPQ}$ even if $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, see Fig. 1.6c for an example. Instead we show an analogous result for when the conditions of Theorem 1.3.1(ii) are met, $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, and when α is sufficiently large, it follows that $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$. The proof can be found in Appendix A.4.

Corollary 1.3.1. *Assume customers are α -mixed and draw their valuations from an exchangeable distribution. Suppose DP earns more than SP and let $\gamma > 0$ denote the gap, i.e., $\mathcal{R}_{DP} = (1 + \gamma)\mathcal{R}_{SP}$.*

(i) If $\alpha = 1$, then

$$\mathcal{R}_{OPQ} > \frac{1 + \frac{N}{N-1}\gamma}{1 + \gamma} \mathcal{R}_{DP}.$$

(ii) If the valuations are supported on two points and $\alpha \geq 1 - \frac{\gamma}{\gamma N + N - 1}$, then

$$\mathcal{R}_{OPQ} > \alpha \left(1 + \frac{\gamma}{(N-1)(1+\gamma)} \right) \mathcal{R}_{DP}.$$

Proof of Theorem 1.3.1.

Proof of Theorem 1.3.1. We will consider the two cases separately.

Case (i): Let $\alpha = 1$, F be the exchangeable distribution over N items, and w.l.o.g. let $p_1 \geq p_2 \geq \dots \geq p_N$ be the optimal prices corresponding to \mathcal{R}_{DP} . For ease of exposition we assume F is continuous and ignore ties, although the same argument follows when F is not continuous and one carefully considers the tie-breaking procedure. Let Σ be the set of permutations $\sigma : [N] \rightarrow [N]$, and $\sigma(i)$ the mapping of index i under σ . For every $\sigma \in \Sigma$, define the event $E_\sigma := \{V_{\sigma(1)} \geq V_{\sigma(2)} \geq \dots \geq V_{\sigma(N)}\}$. Note that each $\{E_\sigma\}_{\sigma \in \Sigma}$ is equally likely by exchangeability. We define $q_{i|\sigma}$ to be the probability of a customer purchasing i under the DP strategy (p_1, \dots, p_N) conditioned on the event E_σ . We define $Rev(p_1, \dots, p_N|\sigma)$ to be the expected revenue of the DP strategy conditioned on the event E_σ , i.e.,

$$Rev(p_1, \dots, p_N|\sigma) = \sum_{i=1}^N p_i q_{i|\sigma}.$$

Define σ^* such that $Rev(p_1, \dots, p_N|\sigma^*) \geq Rev(p_1, \dots, p_N|\sigma)$ over all $\sigma \in \Sigma$, i.e., E_{σ^*} is the event that leads to the most revenue. This implies that $Rev(p_1, \dots, p_N|\sigma^*) \geq \mathcal{R}_{DP}$.

Now consider an opaque selling strategy OPQ that uses prices $p^i = p_{\sigma^*(i)}$. (Note that p^1 is the price of the items.) We shall show that this opaque strategy has expected

revenue of at least $Rev(\vec{p}|\sigma^*)$. Under our opaque strategy, we call the probability of a customer buying an opaque product of size i to be q^i and the probability of a customer buying an item to be q^1 . We let $V^{(i)}$ be the i^{th} order statistic such that $V^{(1)} = \max_i\{V_i\}$ and $V^{(N)} = \min_i\{V_i\}$. We now show that $q^i = q_{\sigma^*(i)|\sigma^*}$ for all i by

$$\begin{aligned}
q^i &= \mathbb{P}\left(\max_{S, |S|=i} \{V^S - p^i\} \geq \max_{j \neq i, |S'|=j} \{V^{S'} - p^j, 0\}\right) \\
&= \mathbb{P}\left(V^{(i)} - p^i \geq \max_{j \neq i} \{V^{(j)} - p^j, 0\}\right) \\
&= \mathbb{P}\left(V^{(i)} - p^i \geq \max_{j \neq i} \{V^{(j)} - p^j, 0\} | E_{\sigma^*}\right) \\
&= \mathbb{P}\left(V_{\sigma^*(i)} - p_{\sigma^*(i)} \geq \max_{j \neq i} \{V_{\sigma^*(j)} - p_{\sigma^*(j)}, 0\} | E_{\sigma^*}\right) \\
&= q_{\sigma^*(i)|\sigma^*}.
\end{aligned}$$

The first equality follows from the definition of OPQ strategies and q^i . The second equality follows from noting that a customer only needs to consider the best opaque product of each possible size $i = 2, \dots, N$ and the best item. The best opaque product of size i has a valuation of the minimum of the top i valuations, which is the i 'th order statistic. The third equality follows from the fact that the valuations are exchangeable, and thus an event on the order statistics is independent of E_σ for all $\sigma \in \Sigma$. The fourth equality follows from our pricing rule and the definition of σ^* . The last equality follows from the definition of $q_{i|\sigma^*}$. Combining our findings yields

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{OPQ}(p^1, \dots, p^N) = \sum_{i=1}^N p^i q^i = \sum_{i=1}^N p_{\sigma^*(i)} q_{\sigma^*(i)|\sigma^*} = Rev(p_1, \dots, p_N | \sigma^*) \geq \mathcal{R}_{DP}. \quad \square$$

Case (ii): Fix a distribution F supported on two points $\{a, b\}$ where $a < b$. Note for distributions, the optimal discriminatory pricing uses prices $\vec{p} = (a, a, \dots, a)$, (b, b, \dots, b) or a mixed pricing where exactly one price (since F is exchangeable it doesn't matter which price) is low (a, b, b, \dots, b) . If either (a, a, \dots, a) or (b, b, \dots, b)

are the the optimal discriminatory pricing given F , then $\mathcal{R}_{SP} = \mathcal{R}_{DP}$ and the claim follows automatically. Suppose $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, then the optimal pricing is the mixed strategy and under a mixed pricing, a discriminatory selling strategy always sells an item. Further we will restrict ourselves to opaque pricings where $p^N = a$, and thus always sell the item. Since the item is always sold in both strategies, we may normalize the support of F to $\{1, 1 + \delta\}$ without changing the ratio $\frac{\mathcal{R}_{DP}}{\mathcal{R}_{OPQ}}$. Now let U be a random variable representing the number of valuations that are equal to $1 + \delta$. When $U = 0$, DP earns revenue of 1. When $U = i \geq 1$, DP earns revenue of $1 + \delta$ with probability $\frac{\binom{N-1}{i}}{\binom{N}{i}} = \frac{N-i}{N}$ and 1 otherwise. (The customer buys the cheap item when they value it at $1 + \delta$.) Then for $i \geq 1$,

$$E[\mathcal{R}_{DP}|U = i] = 1 + \frac{N-i}{N}\delta. \quad (1.1)$$

Consider the following opaque pricing where for $i \in \mathcal{N}$ we let $p^i = 1 + \frac{N-i}{N}\delta$. When $U = 0$, the customer buys the opaque product of size N at price 1, paying the same in the corresponding case in DP. When $U = i \geq 1$, we claim that regardless of whether the customer is pessimistic or risk-neutral, they will purchase an opaque product of size i (or item if $i = 1$) earning revenue $1 + \frac{N-i}{N}\delta$, which is the same revenue in the corresponding case in DP and therefore would complete the proof. First suppose the customer is pessimistic, then when $U = i$ the customer values the size i product as $1 + \delta$ and garners utility $\frac{i}{N}\delta$. For $j < i$, the customer values the opaque product the same but has to pay a higher price, while for $j > i$ the customer values the opaque product at 1 and does not buy. Thus a pessimistic customer yields revenue $1 + \frac{N-i}{N}\delta$ when $U = i$.

When the customer is risk-neutral and $U = i$, they again value the size i product as $1 + \delta$ and garners utility $\frac{i}{N}\delta$ for purchasing it. Products of size $j < i$ have the same valuation, but at a higher price, and thus offer less utility. For products of size

$j > i$, the utility of the size j opaque product is

$$\frac{i(1 + \delta) + (j - i) \cdot 1}{j} - \left(1 + \frac{N - j}{N}\delta\right) = \left(\frac{i}{j} - \frac{N - j}{N}\right)\delta$$

which is strictly less than $\frac{i}{N}\delta$. Finally, the above expression also shows that the utilities of the opaque products of size i and N are the same, in which case the customer buys i (since we have assumed w.l.o.g. that ties are broken in favor of the more expensive option). Thus both pessimistic and risk-neutral customers have the same purchase behavior under this opaque pricing, and yield the same expected revenue as \mathcal{R}_{DP} . \square

In the proof of Theorem 1.3.1(i), one natural thought is to view opaque selling with pessimistic customers as discriminatory pricing where the ordering of the item valuations is known a priori to the seller. That is, the valuations for the best item and opaque products take on exactly the valuations of the original items. It is then tempting to assume that OPQ is trivially more profitable than DP, where the ordering of the item valuations is not known to the seller and hence “less information” is available. However, this false argument would easily extend to other settings where our result does not hold (see Example 1.3.2) and thus the argument is invalid. The extra information offered by OPQ comes with an additional constraint: the highest valued item is sold at the highest price, the second highest valued item is sold at the second highest price, and so on, which need not be optimal.

Benchmarking against Single Pricing

In this section, we seek to quantify the potential gains that opaque selling offers over a simple single pricing strategy. This question has also been studied in the context of discriminatory pricing. For example, when valuations are drawn i.i.d., Chawla et al. [30] shows that discriminatory pricing can earn at most $2 - \frac{1}{N}$ more than single price

strategies, and Dutting and Klimm [49] shows that this bound is tight. Interestingly, in the same setting of i.i.d. valuations, opaque selling can also earn up to a constant factor of single pricing. In Theorem 1.3.2, we describe this upper bound as a function of α and N . A direct consequence of this theorem is that when valuations are i.i.d., OPQ and DP are always within a constant factor of each other. We provide the proof in Section 1.3.

Theorem 1.3.2 (Revenue Upper Bound when Valuations are I.I.D.). *Assume customers are α -mixed and their item valuations are i.i.d. Then,*

$$\mathcal{R}_{OPQ} \leq \left(3 + (1 - \alpha) \left(1 - \frac{2}{N} \right) \right) \mathcal{R}_{SP}.$$

In the more general case of exchangeable distributions, no results comparing DP to SP are available to the best of our knowledge. In Theorem 1.3.3, we show that DP earns at most $1 + \log(N)$ more than SP, while OPQ earns at most N times more than SP. This implies that OPQ can earn up to (order of) $\frac{N}{1 + \log(N)}$ more revenue than DP, which we also show is indeed possible in Theorem 1.3.3 and Example A.3.1. We defer the proof to Appendix A.4

Theorem 1.3.3 (Revenue Upper Bound when Distribution is Exchangeable). *Assume customers are α -mixed and draw their valuations from an exchangeable distribution. Then, (i) $\mathcal{R}_{DP} \leq (1 + \log(N)) \mathcal{R}_{SP}$, (ii) $\mathcal{R}_{OPQ} \leq N \mathcal{R}_{SP}$, and (iii) there exists a distribution F such that $\mathcal{R}_{OPQ} \geq \frac{\alpha}{2} \frac{N}{1 + \log(N)} \mathcal{R}_{DP}$.*

Proof of Theorem 1.3.2

We divide the proof of Theorem 1.3.2 into two lemmas. Lemma 1.3.1 states that when customers are purely pessimistic $\mathcal{R}_{OPQ} \leq 3 \mathcal{R}_{SP}$. Lemma 1.3.3 states that when customers are purely risk-neutral $\mathcal{R}_{OPQ} \leq \left(4 - \frac{2}{N} \right) \mathcal{R}_{SP}$. To obtain Theorem 1.3.2, we relax OPQ to observe X_α and price pessimistic and risk-neutral customers separately.

Using Lemmas 1.3.1 and 1.3.3, we get that $\mathcal{R}_{OPQ} \leq (\alpha \cdot 3 + (1 - \alpha)(4 - \frac{2}{N})) \mathcal{R}_{SP}$ which is the desired result.

Lemma 1.3.1. *Assume all customers are pessimistic. Then when item valuations are i.i.d.,*

$$\mathcal{R}_{OPQ} \leq 3\mathcal{R}_{SP}.$$

The proof of Lemma 1.3.1 below relies on connecting the revenue generated by OPQ to a Myerson auction, and makes use of the following lemma.

Lemma 1.3.2 (Chawla et al. [31] Theorem 8). *Let $\mathcal{R}^{\mathcal{M}}$ be the expected revenue from the Myerson auction for one item, run on N bidders with i.i.d. valuations. Then*

$$\mathcal{R}^{\mathcal{M}} \leq 2\mathcal{R}_{SP}.$$

Armed with this lemma we can now prove Lemma 1.3.1.

Proof of Lemma 1.3.1. Let (p, \vec{p}) denote the prices of an optimal OPQ strategy under F , where p is the price of items and \vec{p} are the prices of the opaque products. The proof follows by separately bounding revenue from items priced at p and the revenue from opaque products. Let $V^{(k)}$ be the k^{th} order statistic (counting so that $V^{(1)} = \max_i \{V_i\}$), and note that the highest valuation a customer has for opaque

products of size k is just $V^{(k)}$. Then,

$$\begin{aligned}
\mathcal{R}_{OPQ} &= p\mathbb{P}(V^{(1)} - p \geq \max_{k=2,\dots,N} \{V^{(k)} - p^k, 0\}) + \sum_{k=2}^N p^k \mathbb{P}(\text{buys opaque product of size } k) \\
&\leq p\mathbb{P}(V^{(1)} - p \geq 0) + \sum_{k=2}^N p^k \mathbb{P}(\text{buys opaque product of size } k) \\
&\leq p\mathbb{P}(V^{(1)} - p \geq 0) + E[V^{(2)}] \\
&\leq \mathcal{R}_{SP} + E[V^{(2)}] \\
&\leq \mathcal{R}_{SP} + \mathcal{R}^{\mathcal{M}} \\
&\leq 3\mathcal{R}_{SP}.
\end{aligned}$$

The equality follows from the definitions of \mathcal{R}_{OPQ} , p , and \vec{p} . The first inequality follows from non-negativity of $\max_{k=2,\dots,N} \{V^{(k)} - p^k, 0\}$. The second inequality follows from realizing that the highest valued opaque product is valued at $V^{(2)}$, and thus customers pay at most $V^{(2)}$ when buying an opaque product. The third inequality follows from the optimality of \mathcal{R}_{SP} . The fourth inequality follows from the fact that $E[V^{(2)}]$ is the revenue of a second price auction, which is dominated by the Myerson auction. The final inequality follows from Lemma 1.3.2. \square

We now consider the case of risk-neutral customers in Lemma 1.3.3.

Lemma 1.3.3. *Assume all customers are risk-neutral. When item valuations are i.i.d., then*

$$\mathcal{R}_{OPQ} \leq \left(4 - \frac{2}{N}\right) \mathcal{R}_{SP}.$$

The proof of Lemma 1.3.1 bounds the revenue from exponentially many opaque products by the highest valuation any opaque product could receive from a pessimistic customer. We noted that the highest valuation for an opaque product is bounded by the expected value of a second price auction, which allowed us to apply Lemma 1.3.2. Such an argument fails for risk-neutral customers since valuations for opaque products

can be higher than $V^{(2)}$, the second order statistic of V . To circumvent this difficulty, we recast opaque selling with risk-neutral customers in the language of lotteries.

Definition 1.3.1. *A lottery over N items denoted by (p, \vec{q}) consists of a price p and probabilities q_i for receiving each item i , s.t. $\sum_{i=1}^N q_i \leq 1$.*

A customer with valuation vector \vec{v} values a lottery (p, \vec{q}) as $\sum_{i=1}^n v_i q_i - p$. Note that selling lotteries can simulate deterministic item pricing by defining N lotteries where lottery $l_i = (p, e_i)$, where e_i is the i^{th} unit vector. An opaque product over a set S can be cast as a lottery with price $p^{|S|}$ and allocation probabilities $q_i = \frac{1}{|S|}$ for each $i \in S$ and $q_i = 0$ for each $i \notin S$. We call a collection of offered lotteries a *lottery pricing*, denoted by \mathcal{L} . Using this framework, we can prove that OPQ can obtain at most $4 - \frac{2}{N}$ times more revenue than SP. The proof can be found in Appendix A.4, draws on lottery pricing results of Chawla et al. [32], who proved an upper bound of 4 in their setting.

1.4 The Power of One Opaque Product

In this section, we study the revenue gained by using a strategy with a single opaque product (1OPQ), where the seller offers all N items at a single price alongside a single opaque product corresponding to the set \mathcal{N} . 1OPQ represents the easiest use-case for opaque selling, simply offering one opaque option made up of all N items. Fig. 1.1 shows an example of 1OPQ for staplers on Amazon.com.

We note that since the 1OPQ strategy only offers two prices, then a comparison to discriminatory pricing which offers N prices becomes significantly more challenging. Nevertheless, we show in Section 1.4 that 1OPQ guarantees 71.9% of the revenue of DP in the special case of two-point distributions. In comparison to single pricing, we show that 1OPQ can earn at most a factor of $(2 - (1 - \alpha)\frac{1}{N})$ more than SP in

Section 1.4. When $N = 2$, our upper bounds are tight and the revenue increase can be larger than that of DP.

Benchmarking against Discriminatory Pricing

In Theorem 1.4.1 we show that 1OPQ guarantees at least 71.9% of the revenue that DP provides when the distribution is exchangeable and valuations are supported on two points (low or high). As previously mentioned, such distributions are a natural and well-studied model of customers with binary preferences, and may be used to approximate bimodal distributions. Further, as seen in Example A.3.2 and Chawla et al. [30], two point distributions represent natural best cases for price discrimination for both 1OPQ and DP strategies. We emphasize that Theorem 1.4.1 is a strict improvement on the approximation possible by \mathcal{R}_{SP} , which is $0.50\mathcal{R}_{DP}$ in this setting. Specifically, Chawla et al. [30] give a two point distribution such that when scaling the number of items N , $\lim_{N \rightarrow \infty} \frac{\mathcal{R}_{SP}}{\mathcal{R}_{DP}} = .5$.

Theorem 1.4.1 (When 1OPQ Approximates DP). *If customers are α -mixed and draw their valuations from an exchangeable distribution supported on two values, then*

$$\mathcal{R}_{1OPQ} \geq .719\mathcal{R}_{DP}.$$

Our proof follows from observing that when the probability of customers having high valuations is large, a single pricing strategy is a good approximation. Otherwise, if the probability of customers having high valuations is small, we show that augmenting single price strategies with a single opaque product is a good approximation of the optimal discriminatory pricing. We defer the complete proof to Appendix A.4.

Benchmarking against Single Pricing

In this section, we show that the addition of a single opaque product over the set \mathcal{N} can increase the revenue by at most $(2 - (1 - \alpha)\frac{1}{N})$. Although our bound holds under all exchangeable distributions, Examples A.3.2 and A.3.3 shows that our analysis is tight in the special cases of $\alpha = 0$ and $\alpha = 1$, even when customers are restricted to have i.i.d. valuations. We defer the proof to Appendix A.4.

Theorem 1.4.2 (Revenue Upper Bounds for 1OPQ). *Assume customers are α -mixed and draw their valuations from any distribution. Then,*

$$\mathcal{R}_{1OPQ} \leq \left(2 - (1 - \alpha)\frac{1}{N}\right) \mathcal{R}_{SP}$$

Further, when $\alpha = 0$ or $\alpha = 1$ there exists an i.i.d valuation distribution such that the bounds are tight.

Theorem 1.4.2 fully describes the possible revenue increase a seller could hope to garner using a single opaque product. It is of interest to note that when $N = 2$ and $\alpha = 1$, Theorem 1.4.2 implies the existence of a valuation distribution such that $\mathcal{R}_{1OPQ} = 2\mathcal{R}_{SP}$. However by Theorem 1.3.3(i), $\mathcal{R}_{DP} \leq (1 + \log(2))\mathcal{R}_{SP}$ for any distribution. Together these results show that 1OPQ can sometimes achieve higher revenue lifts than DP.

1.5 Numerical Experiments

In this section we conduct numerical experiments to demonstrate the possible relationships between $\mathcal{R}_{SP}, \mathcal{R}_{1OPQ}, \mathcal{R}_{DP}, \mathcal{R}_{OPQ}$ for various valuation distributions. To perform the experiments, we must solve for the optimal prices for any of these strategies. That is given the distribution F , and α , we must solve for the price vector that maximizes revenue. However, in general solving for the optimal prices in multi-item

settings is quite difficult. Even in the special case when the valuations are i.i.d., solving for the optimal discriminatory pricing strategy is NP-Hard Chen et al. [37].

Although not the focus of this work, in Section 1.5 we address this issue by developing a simple enumerative algorithm which is computationally efficient in the special case when the support of the valuations is discrete and the number of items is not large. Note that a simple brute force search over the support is not sufficient, as optimal prices do not necessarily lie on the support (Chawla et al. [30], Chen et al. [35, 37]). Given any distribution F , our approach is to first discretize the distribution and then run Algorithm 1. When the number of items is large, one can employ a standard MIP approach along the lines of Hanson and Martin [66].

We emphasize that carefully discretizing the support and then solving still yields near-optimal solutions for the true underlying distribution. Indeed, Hartline and Koltun [70] show that when valuations are supported on $[l, h]$, generating $\log_{1+\epsilon} \left(\frac{h}{l}\right)$ discrete points and solving obtains prices that garner revenue within a factor of $1 + \epsilon$ of the optimal revenue. In Section 1.5, we conduct our experiments by using the discretization approach alongside our enumerative algorithm.

An Enumerative Algorithm for Finding Optimal Pricings

In this section, we describe an algorithm for finding the optimal pricing in the special case when the support of the distribution is discrete. (As mentioned previously, we assume that we have approximated the original distribution by a discrete distribution.) When the number of items is small, this algorithm is relatively efficient. Specifically, we show that if N is assumed to be (a small) constant, then the optimal prices for any strategy (SP, DP, OPQ, or 1OPQ) can be found in time that is polynomial in the size of the support of the valuation distribution.

We let m denote the number of points (valuation vectors) in the support of F . Each point j corresponds to a customer type with a valuation vector $\vec{v}_j =$

$(v_{j,1}, v_{j,2}, \dots, v_{j,N})$. When referring to DP, $v_{j,i}$ denotes type j 's valuation for item i . When referring to OPQ, $v_{j,i}$ denotes type j 's valuation for the opaque product of size i .

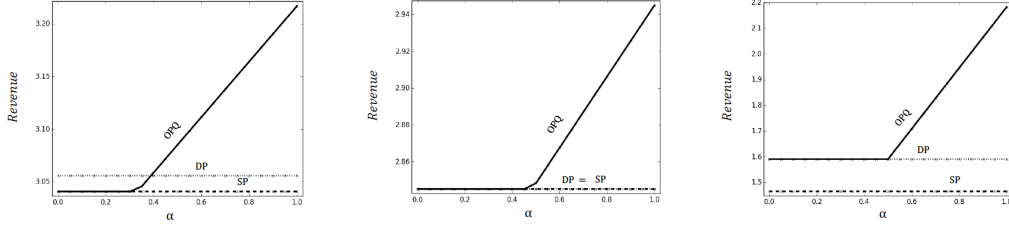
Note that opaque valuation vectors can easily be generated given a discretized distribution by computing the opaque valuations for pessimistic and risk-neutral customers directly. The type vector can also be made to correspond to the type vector for SP by replacing each vector in the DP case with $(\max_i v_{j,i})$, and 1OPQ by replacing each vector in the OPQ case with $(v_{j,1}, v_{j,N})$. In Theorem 1.5.1 we show that the optimal prices can be found in time $O(m^N)$. The idea of Algorithm 1 is to identify a set of $(m+1)^N$ candidate prices, which is guaranteed to contain the optimal price. The algorithm then enumerates over the set of candidate prices and returns the price that yields the highest revenue. We defer the details to Appendix A.4.

Theorem 1.5.1 (Algorithm for Computing Optimal Prices). *Let F be an exchangeable distribution over m customer types. Then both the optimal opaque pricing and discriminatory pricing can be computed in $O(m^N)$ time by Algorithm 1.*

Computational Results

In this section we conduct numerical experiments on three typical valuation distributions which bear out the relationships between \mathcal{R}_{SP} , \mathcal{R}_{DP} , \mathcal{R}_{OPQ} , and \mathcal{R}_{1OPQ} that we have studied in the previous sections. We shall assume item valuations are drawn i.i.d. from the following distributions: (i) a triangular distribution supported on $[1, 7]$ with mode 3, (ii) a normal distribution with mean 3 and standard deviation 2 truncated on $[1, 7]$, and (iii) a Bernoulli distribution supported on $\{1, 7\}$ with probability of a 7 being $1/9$. In order to apply Theorem 1.5.1, we discretize these distributions by rounding valuations to their nearest integer value. We compute the revenue of SP, DP, OPQ, and 1OPQ for every value of α from 0 to 1 in increments of 0.05. Fig. 1.6

displays our results when $N = 2$ (in which case OPQ is equivalent to 1OPQ) and Fig. 1.7 displays our results when $N = 3$.

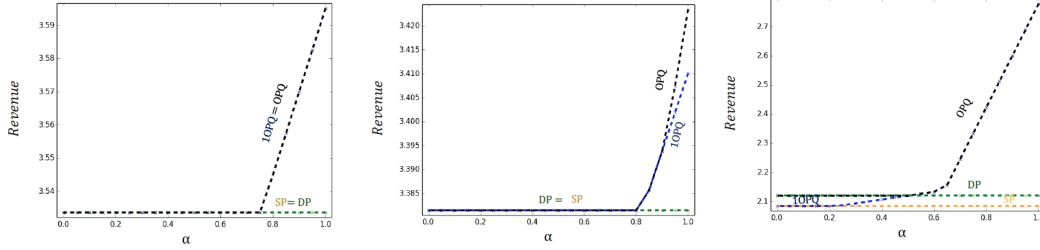


(a) Triangular distribution. (b) Normal distribution. (c) Bernoulli distribution.

Figure 1.6: Illustrates the relationship between \mathcal{R}_{SP} (dashed line), \mathcal{R}_{DP} (dotted line), and \mathcal{R}_{OPQ} (solid line) as the proportion of pessimistic customers increases.

Each of the three distributions we study result in fundamentally different behaviors. In Fig. 1.6(a) we note that $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, and further when $\alpha < 0.3$, $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$. However as α increases towards one, the relationship between \mathcal{R}_{DP} and \mathcal{R}_{OPQ} reverses. When α is close to 1, \mathcal{R}_{OPQ} significantly outperforms \mathcal{R}_{DP} , garnering up to approximately 5% more revenue. In Fig. 1.6(b) we note that $\mathcal{R}_{SP} = \mathcal{R}_{DP}$, meaning that discriminatory pricing alone does not add value over a single price. However, OPQ can earn strictly more revenue than either strategy when $\alpha > 0.5$. Finally in Fig. 1.6(c), $\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}$ for any value of α , which is known directly from Theorem 1.3.1(ii). The gap is positive and increasing when $\alpha > 0.5$, which is implied by Corollary 1.3.1. Interestingly, for all three distributions \mathcal{R}_{OPQ} is a non-decreasing function in α . This is counterintuitive: as the number of pessimistic customers increase, more customers have lower values for the opaque products but the overall revenue from OPQ increases. This suggests that the revenue non-monotonicity noted in Section 1.3 is quite pervasive.

In Fig. 1.7(a), we see that that lifting the problem from $N = 2$ to $N = 3$ collapses the revenue gap between \mathcal{R}_{SP} and \mathcal{R}_{DP} , but does not diminish the impact of opaque products. Further we note that a single opaque product performs just as well as the general opaque strategy does. In Fig. 1.7(b), we observe that when $\alpha > 0.8$, 1OPQ



(a) Triangular distribution. (b) Normal distribution. (c) Bernoulli distribution.

Figure 1.7: Illustrates the relationship between \mathcal{R}_{SP} , \mathcal{R}_{DP} , \mathcal{R}_{1OPQ} and \mathcal{R}_{OPQ} as the proportion of pessimistic customers increases.

and OPQ can outearn DP. As α approaches 1, eventually there is a revenue gap between 1OPQ and OPQ. Finally in Fig. 1.7(c), we see that when $\alpha < 0.5$, DP and OPQ are equivalent and outperform 1OPQ. When $\alpha > 0.5$, 1OPQ and OPQ become equivalent and outperform DP. We believe these experiments demonstrate a wide range of behavior, but generally the OPQ and 1OPQ strategies tend to outperform DP in almost all cases, and tend to improve as α increases.

1.6 Conclusion

In this paper, we studied opaque selling strategies in the context of selling horizontally differentiated items to unit-demand, utility-maximizing customers. We considered mixtures of two practical models of customer behavior corresponding to pessimistic and risk-neutral customers, motivated by the customer’s lack of knowledge about how opaque products are allocated by the seller. When the valuation distribution is exchangeable and either customers are pessimistic or have binary preferences, we showed that opaque selling dominates discriminatory pricing. We also explicitly quantified the best possible revenue lift from using opaque products, which can be significantly higher than discriminatory pricing. Finally, we considered the practical case where only one grand opaque product is offered, and offered theoretical and numerical evidence of the strength of this simplified strategy.

We believe our results provide strong theoretical motivation for using opaque products as a vehicle for price discrimination, especially in online sales channels. Since our opaque model imposes a single price for opaque products of the same size, it is impartial to customers with particular preferences. It is also particularly advantageous in situations where discriminatory pricing could be effective, but disallowed due to business constraints and poor customer perception. It would be interesting for future research to consider the impact of competition and finite inventory constraints on opaque selling, as well as behavioral studies for how consumers value opaque products in various markets.

Chapter 2

The Value of Personalized Pricing

In this chapter we study personalized pricing strategies. The contents of this section are joint work with *Adam Elmachtoub* and *Vishal Gupta*.

Increased availability of high-quality customer information has fueled interest in personalized pricing strategies, i.e., strategies that predict an individual customer's valuation for a product and then offer a customized price tailored to that customer. While the appeal of personalized pricing is clear, it may also incur large costs in the form of market research, investment in information technology and analytics expertise, and branding risks. In light of these tradeoffs, our work studies the value of idealized personalized pricing over a spectrum of pricing strategies varying in pricing flexibility and prediction model accuracy.

We first provide tight, closed-form upper and lower bounds on the ratio between the profits of an idealized personalized pricing strategy and a single price strategy. These bounds depend on simple statistics and/or shape assumptions of the valuation distribution and shed light on the types of markets for which personalized pricing has the most potential. Next, we consider two stylized price discrimination strategies that isolate the key assumptions underlying idealized personalized pricing: (i) a feature-based pricing strategy, where the firm can charge a continuum of prices, but does not know the customers valuations precisely (ii) a k -market segmentation strategy where the firm knows all customer valuations precisely but can only charge customers one of k prices. For each strategy we bound the ratio of idealized personalized pricing profits to the profits of that strategy. We then synthesize these results to study a

more realistic personalization strategy in which the seller neither knows customer valuations precisely nor is able to offer a continuum of prices. These bounds quantify the value of the operational capability of charging distinct prices and the value of additional predictive accuracy, respectively.

Finally, we generalize our work in two directions: (i) we show how to extend bounds on the value of personalized pricing to stronger bounds on the ratio of feature-based pricing and a single price strategy and (ii) how to obtain bounds on the value of personalized pricing that depend on arbitrary moments via infinite dimensional linear programming duality.

2.1 Introduction

Over the last decade, increased availability of high-quality customer information has fueled interest in personalized pricing strategies. At a high-level, these strategies combine customer data with machine learning and optimization tools to predict an individual customer’s willingness to pay and then customize a price for that customer. This customized price can be delivered as a discount via a mobile application or other channel.

The appeal of personalized pricing is clear – If a seller could accurately predict individual customer valuations, then it could (in principle) charge each customer exactly their valuation, increasing profits and market penetration. Given this appeal, grocery chains [40], department stores [47], airlines [101], and many other industries [83] have begun experimenting with personalized pricing. Moreover, within the operations community, there has been a surge in research on how to practically and effectively implement personalized pricing strategies (e.g., Aydin and Ziya [7], Phillips [87], Bernstein et al. [18], Chen et al. [36], Ban and Keskin [13]).

Unfortunately, implementing any form of price discrimination, including person-

alized pricing, may be costly and/or difficult. A firm would need to engage in price experimentation and market research, invest in information systems to store customer data, and build analytics expertise to transform these data into a personalized pricing strategy (see Arora et al. [6] for an extensive discussion). Moreover, price discrimination tactics involve serious branding risks and potential customer ill-will, and, in some markets, may be of questionable legality. Finally, personalized pricing may impact competitors' [108] and manufacturers' [78] behavior.

In light of these tradeoffs, in this work we complement the existing operations literature on *how* to implement personalized pricing by quantifying *when* personalized pricing offers significant value. Specifically, for a single-product monopolist, we bound the profit ratio between *idealized personalized pricing (PP)*, i.e., charging each customer exactly their willingness to pay, and a spectrum of various simpler pricing strategies. The spectrum of strategies vary on the degree of pricing flexibility as well as prediction model accuracy. Thus, these bounds can guide managers in assessing the potential upside of the above tradeoffs, and provide a fundamental understanding of the value of offering more prices and of the value of reducing prediction error.

With full-information about the customer valuation distribution, computing the exact ratio between idealized personalized pricing over simpler pricing strategies is straightforward; there is no need for bounding. However, in our opinion, a firm not currently engaging in personalized pricing is unlikely to know the full valuation distribution. Indeed, it is not necessary to learn this distribution to price effectively [21, 20] and learning it may be difficult since real-world distributions are typically complex and irregular (see, e.g., Celis et al. [28] for a discussion in an auction setting).

Consequently, we focus instead on parametric bounds that depend on a few statistics of the valuation distribution. On the one hand, we believe these statistics are more easily estimated by a seller not currently engaging in personalized pricing than the full valuation distribution. On the other hand, and perhaps more importantly,

parametric bounds based on these statistics provide structural insights into the types of markets where the value of personalized pricing is potentially large. In particular, we leverage these structural insights to disentangle the contributions from increased operational flexibility (offering many distinct prices) and improved prediction accuracy (gathering additional data) in personalized pricing strategies.

More specifically, in the first part of the paper, we bound the profit ratio between idealized personalized pricing and posting a *single price* (SP) for all customers. We call this ratio the *value of personalized pricing over single-pricing*. Notice that idealized personalized pricing as we define it is often called first-degree price discrimination in the economics literature, and observe that it upper bounds the profit of any other price discrimination strategy. Thus, the value of personalized pricing over single-pricing also upper bounds the potential gains of any other price discrimination strategy over single pricing.

We prove bounds that are tight, closed-form and depend on three unitless statistics of the valuation distribution: (i) the *scale*, which is the ratio of the upper bound of the support to the mean, (ii) the *margin*, which we define as the margin of a unit sold at a price equal to the mean valuation, and (iii) the *coefficient of deviation*, which is the mean absolute deviation over twice the mean. Knowing these three quantities is equivalent to knowing the mean, support, and mean absolute deviation of the distribution. Our bounds are tight in the sense that we give an explicit valuation distribution for which the value of personalized pricing over single-pricing matches the bound. The precise form of the tight distribution depends on the relevant parameters, but consists of a mixture of Pareto and two-point distributions. These results generalize folklore results that the Pareto distribution (a.k.a. “equal-revenue” distribution) represents the worst-case for single-pricing. Perhaps surprisingly, we also find that our bound is maximal for intermediate values of the coefficient of deviation and approaches one as the coefficient deviation increases with all other parameters fixed.

We complement our upper bounds with novel lower bounds depending on the coefficient of deviation and a mild shape assumption, together our bounds yield strong conditions when personalized pricing is necessary or superfluous.

Of course, idealized personalized pricing is not achievable in practice. It hinges on two assumptions: First, the monopolist has the ability to charge a potentially distinct price to each customer. Second, the monopolist is omniscient and can perfectly predict each customer’s valuation. In the second part of this paper, we study price discrimination strategies that relax these two assumptions and more closely model personalized pricing strategies used in practice. To this end, we first compute the value of personalized pricing over two stylized price-discrimination strategies: k -market segmentation and feature-based pricing.

In the *k-market segmentation* (kP) strategy, we assume the monopolist is still omniscient, but can charge at most k distinct prices, relaxing the assumption of a continuum of prices. Thus, the value of personalized pricing over k -market segmentation quantifies the value of the operational capability of charging a continuum of prices, which is equivalent to the case where $k \rightarrow \infty$. Under a mild assumption, we show that this value is at most $1 + C\frac{1}{k}$, where C is an explicit constant depending on distributional parameters. We prove theoretically that this worst-case dependence on k is tight and provide numerical evidence that it is in fact typical of many distributions. This analysis yields a natural rule of thumb; to half the gap to the ideal personalized pricing profits, one needs to double the number of prices offered.

By contrast, in the *feature-based pricing* (XP) strategy, we assume the monopolist can in principle offer a continuum of prices, but is no longer omniscient. Rather, she observes a feature vector (sometimes called a context) for each customer which she can use to (imperfectly) predict the customer’s valuation. Thus, the value of personalized pricing over feature-based pricing quantifies the value of additional information, i.e., a richer set of features that would enable perfect prediction. Leveraging our earlier

results, we prove that this value is bounded by an explicit factor that depends on the coefficient of deviation of the error in the valuation prediction model. Thus, our bound quantifies the degree of prediction accuracy necessary to guarantee a certain percentage of profits. Again, we provide numerical evidence suggesting our worst-case analysis is qualitatively typical of many valuation distributions. Our result yields another natural rule of thumb; to half the gap to the ideal personalized pricing profits, one needs to quadruple the prediction accuracy.

We use the above results on these stylized pricing strategies as building blocks to study a more realistic *feature-based market segmentation* (kXP) strategy. In this strategy, we assume the monopolist is neither omniscient nor operationally able to offer a continuum of prices. Rather, as in the feature-based pricing strategy, she observes a feature for each customer. Based on this feature, she then offers the customer one of k prices. Bounding the relative difference between idealized personalized pricing and feature-based market segmentation quantifies the impact of both limited price flexibility and prediction error on the profit of personalized pricing strategies. We believe that feature-based market segmentation also closely resembles data-driven price discrimination strategies commonly used in practice. Under mild assumptions, we show that one can decompose the value of personalized pricing over feature-based market segmentation by separately considering the profit loss from prediction inaccuracy and the profit loss from limited price flexibility on a related, “de-noised” market. These two losses can be analyzed directly using the previously discussed bounds. Importantly, our decomposition is constructive and yields an algorithm for generating a feature-based market segmentation strategy with a provable performance guarantee. In Fig. 2.1 we visualize the relations between all of our various pricing strategies.

Finally, we show how to generalize our work beyond idealized personalized pricing and beyond the coefficient of deviation. Specifically, we prove a novel extension theorem that generalized bounds on the ratio between idealized personalized pricing and a

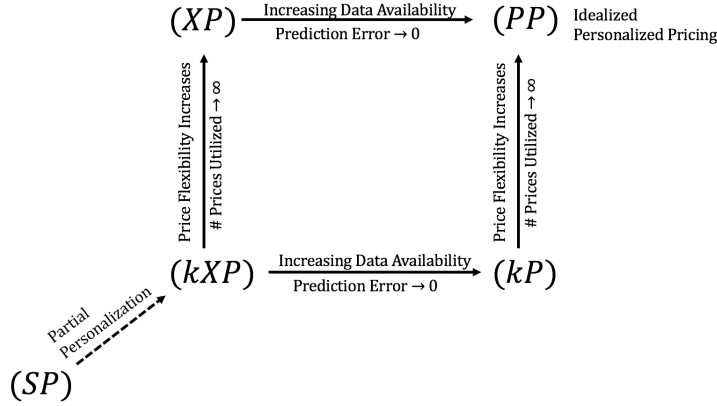


Figure 2.1: We represent our pricing models in terms of their *price flexibility* (y-axis) and *prediction accuracy* (x-axis).

single price strategy to bounds on the ratio between feature-based pricing and a single price strategy. Further, all of our bounds depends on the margin, scale and coefficient of deviation of the valuation distribution. It is possible to generate similar bounds for other statistics. In this vein, we provide an algorithmic procedure to compute an essentially tight bound on the value of personalized pricing over single-pricing given any generalized moment of the valuation distribution, e.g., its variance or geometric mean. The key ideas leverage continuous linear optimization duality and a careful discretization to construct a near-optimal dual feasible solution. The algorithm is provably computationally tractable under mild assumptions on the function defining the generalized moment. These assumptions are satisfied by the usual typical moments encountered in practice. We show our procedure significantly outperforms the best-known bound for the geometric mean, and has a similar behavior as our closed-form bound when using coefficient of variation.

To summarize our contributions:

1. We prove closed-form, tight upper and lower bounds for the value of personalized pricing over single-pricing when the scale, margin, and coefficient of deviation of the valuation distribution are known (cf. Theorem 2.3.1) or when the dis-

tribution is unimodal and left-skew (cf. Theorem 2.3.2). These bounds yield a sharp characterization when personalized pricing is necessary or superfluous.

2. We prove closed-form bounds on the value of personalized pricing over k -market segmentation which are tight in their dependence on k , and describe a distribution-agnostic segmentation procedure that achieves this bound (cf. Theorem 2.4.2). We provide numerical evidence that this worst-case dependence is in fact typical. Thus, the bound quantifies the operational value of being able to charge infinitely many prices over k prices.
3. We further prove closed-form bounds on the value of personalized pricing over feature-based pricing (cf. Theorem 2.4.1). The bound gives an explicit relationship between the accuracy in predicting valuations and value of personalized pricing. Thus, these bound help quantify the value of additional consumer data.
4. We analyze the value of personalized pricing over feature-based market segmentation by synthesizing our results on k -market segmentation and feature-based pricing. We show that the feature-based market segmentation strategy’s profit loss can be bounded by the sum of the profit loss from feature-based pricing and the profit loss of k -market segmentation on a related, “noiseless” market (cf. Theorem 2.4.3). Thus, the decomposition yields a useful, rigorous paradigm for tuning personalized pricing strategies by explicitly identifying the impact on revenue due to limited price flexibility and prediction error.
5. We generalize our above bounds in two ways (i) we provide a novel extension theorem transforms bounds on the ratio between idealized personalized pricing and single price strategies to stronger bounds on the ratio between feature-based pricing and a single price strategy (cf. Theorem 2.5.1) and (ii) we provide a general methodology for computing essentially tight upper bounds on the value of personalized pricing over single-pricing when the scale, margin, and a generalized moment of the valuation distribution other than the coefficient of

deviation are known (cf. Theorem 2.5.2).

Connections to Existing Literature

The study of price discrimination tactics has a long history in economics dating back at least to Robinson [92]. Historically, the economics literature has focused on how various forms of price discrimination affect social welfare (see, e.g., Narasimhan [82], Schmalensee [94], Varian [102], Shih et al. [97] or Bergemann et al. [17], Cowan [45], Xu and Dukes [106] for more recent results). In contrast to these works, we take an operational perspective, focusing on the individual firms relative profits under first-degree price discrimination and other forms of pricing.

That said, we are not the first to study the value of personalized pricing over single pricing. Previous authors have also studied the value of personalized pricing over single pricing under different distributional assumptions. Barlow et al. [14] prove that if the valuation distribution has monotone hazard rates, the value of personalized pricing is at most Euler’s constant $e \approx 2.718$. Similarly, Tamuz [99] prove that if the ratio of the geometric mean over the mean is at least $1 - \delta$, then the value of personalized pricing is at most $(1 - 2^{\frac{4}{3}}\delta^{\frac{1}{3}})^{-1}$. Our single-pricing results differ from these existing results in two critical ways. First, our bounds are tight in the input parameters. Indeed, we show numerically they can be significantly stronger than these existing bounds. Second, since our bounds explicitly depend on simple statistics of the valuation distribution such as the scale and coefficient of deviation, we argue that is easier to use them to assess the effects of various operational decisions. In particular, the parametric dependence on the support and mean absolute deviation of the distribution directly enables us to study personalization strategies which segment markets and leverage features to reduce valuation uncertainty, respectively. It is less clear how to use existing techniques to develop similar results for such strategies.

In Muoz Medina and Vassilvitskii [81] they develop upper bounds that are equiva-

lent to bounds on the value of personalized pricing over single pricing given knowledge of the standard deviation of the valuation distribution. However, the bounds they derive are not tight in contrast to our bounds which are tight for all input parameters we consider. Moreover, Muoz Medina and Vassilvitskii [81] does not consider the roles of price flexibility and prediction accuracy which are central to our paper.

As mentioned above, idealized personalized pricing (first-degree price discrimination) is an idealized strategy. In practice, firms implement some form of third-degree price discrimination. While approaches differ widely, most implicitly or explicitly leverage some form of market segmentation – either segment customers directly or incentivize them to self-segment, and then offer different prices to each segment. Indeed, the operations literature contains many examples of such strategies including intertemporal pricing (Su [98], Besbes and Lobel [19]), opaque selling (Jerath et al. [74], Elmachtoub and Hamilton [51]), rebates/promotions (Chen et al. [38], Cohen et al. [43]), markdown optimization (Caro and Gallien [27], Özer and Zheng [84]), product differentiation (Moorthy [80], Choudhary et al. [39]), dynamic pricing and learning (Cohen et al. [42], Qiang and Bayati [91], Javanmard and Nazerzadeh [73]), and many others.

By contrast, the focus of our work is not on “how to price discriminate” but rather the value of price discrimination. Nonetheless, our bounds on the value of personalized pricing over feature-based market segmentation do provide insight into the above strategies. Since the strategy relaxes the two key assumptions of first-degree price discrimination, our bounds help establish guarantees on the performance of imperfect personalized pricing strategies. Perhaps more importantly, by characterizing the types of valuation distributions for which the value of personalized pricing is high, our bounds highlight the settings in which it is most important that the above strategies perform well and inform their analysis.

Finally, we contrast our work to several recent works that study how to set a

single-price near-optimally given limited distribution information such as the support [41], mean and variance [33, 8], or a neighborhood containing the true valuation distribution [15]. Indeed, these works support our earlier claim that it is not generally necessary to learn the whole valuation distribution in order to price effectively, but are very different in perspective from our work.

2.2 Model and Preliminaries

We consider a profit-maximizing monopolist selling a product with per unit cost c . A random customer's valuation for the product is denoted by the non-negative random variable $V \sim F$. The mean valuation $E[V]$ is denoted by μ . Since it is never profitable to sell to customers with valuations less than c , assume without loss of generality, that $V \geq c$ almost surely. We consider a spectrum of five pricing strategies for the monopolist:

1) Single Pricing (SP): In the single pricing strategy, the monopolist offers the product to all customers at the same price p . Thus, the probability that a customer purchases is given by the complementary cumulative distribution function (cCDF) $\bar{F}(p) := 1 - F(p)$, and the seller's corresponding expected profit is $(p - c)\bar{F}(p)$. Let $\mathcal{R}_{SP}(F, c) := \max_p \{(p - c)\bar{F}(p)\}$ denote the seller's maximal expected profit under single-pricing.

2) Feature-Based Pricing (XP): In the feature-based pricing strategy, the monopolist observes a feature vector \mathbf{X} supported on \mathcal{X} for each customer before offering a price, but does *not* directly observe her valuation V . Based on \mathbf{X} , she offers a customized price $p(\mathbf{X})$, and the customer purchases with probability $\Pr(V \geq p(\mathbf{X}) \mid \mathbf{x})$. Note that unlike k -market segmentation, if \mathbf{X} is continuous, the seller can in principle offer a continuum of prices, one for each possible value of \mathbf{X} . Given a joint distribution $F_{\mathbf{X}V}$ of (\mathbf{X}, V) , let $\mathcal{R}_{XP}(F_{\mathbf{X}V}, c) \equiv \max_{p(\cdot)} E[(p(\mathbf{X}) - c)\mathbb{I}(V \geq p(\mathbf{X}))]$ denote

the optimal profit under feature-based pricing.

3) k -Market Segmentation (kP): In the k -market segmentation strategy, the monopolist partitions the valuation space into $k + 1$ disjoint intervals $[s_0, s_1), [s_1, s_2), \dots, [s_{k-1}, s_k), [s_k, s_{k+1})$ with $s_0 = c$ and assigns distinct prices $p_i \in [s_i, s_{i+1})$ to each segment $i \geq 1$. When a customer with valuation $V \in [s_i, s_{i+1})$ arrives, she is offered the product at price p_i . Given F and c , let $\mathcal{R}_{kP}(F, c, \mathbf{s}, \mathbf{p})$ denote the profit from this strategy with partition \mathbf{s} and prices \mathbf{p} , and let $\mathcal{R}_{kP}(F, c) \equiv \max_{\mathbf{s}, \mathbf{p}} \mathcal{R}_{kP}(F, c, \mathbf{s}, \mathbf{p})$ denote the optimal profit for this strategy.

4) Feature-Based Market Segmentation (kXP): In the feature-based market-segmentation strategy, the monopolist observes a feature \mathbf{X} for each customer, but again does not directly observe her valuation V . Based on \mathbf{X} , she offers one of k prices, $p(\mathbf{X}) \in \{p_i\}_{i=1}^k$, and the customer purchases with probability $\Pr(V \geq p(\mathbf{X}) | \mathbf{X})$. The monopolist's choice of pricing function naturally induces a partition of the market into k segments $\mathcal{X}_i = \{\mathbf{x} \in \mathcal{X} | p(\mathbf{x}) = p_i\}$, and yields expected profit $\sum_{i=1}^k (p_i - c) \Pr(V \geq p_i | \mathbf{X} \in \mathcal{X}_i) \Pr(\mathbf{X} \in \mathcal{X}_i)$. Given a joint distribution $F_{\mathbf{X}V}$ of (\mathbf{X}, V) , let $\mathcal{R}_{kXP}(F_{\mathbf{X}V}, c) \equiv \max_{p_1, \dots, p_k, \mathcal{X}_1, \dots, \mathcal{X}_k} \sum_{i=1}^k (p_i - c) \Pr(V \geq p_i | \mathbf{X} \in \mathcal{X}_i) \Pr(\mathbf{X} \in \mathcal{X}_i)$ denote the optimal profit for this strategy.

5) Idealized Personalized Pricing (PP): In the idealized personalized pricing strategy, the monopolist can potentially offer a different price to each customer and has full knowledge of each customer's valuation. Since $V \geq c$, it is optimal to offer each customer precisely her valuation. Let $\mathcal{R}_{PP}(F, c) := \mu - c$ denote the seller's maximal expected profit under idealized personalized pricing.

By construction, $\mathcal{R}_{SP}(F, c) \leq \mathcal{R}_{kXP}(F_{\mathbf{X}V}, c) \leq \mathcal{R}_{kP}(F, c) \leq \mathcal{R}_{PP}(F, c)$ and $\mathcal{R}_{SP}(F, c) \leq \mathcal{R}_{kXP}(F, c) \leq \mathcal{R}_{XP}(F_{\mathbf{X}V}, c) \leq \mathcal{R}_{PP}(F, c)$. However, in general the ordering between $\mathcal{R}_{kP}(F, c)$ and $\mathcal{R}_{XP}(F_{\mathbf{X}V}, c)$ is instance dependent. Given F and c , we define the *value of personalized pricing over single-pricing* as $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)}$. The

value of personalized pricing over k -market segmentation, feature-based pricing and feature-based market segmentation are each defined similarly. When F , $F_{\mathbf{X}_V}$, and c are clear from context, we sometimes omit them and write, e.g., $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$.

The Lambert- W Function

Many of our closed-form bounds involve $W_{-1}(\cdot)$, the negative branch of the *Lambert- W* function. Although the Lambert- W function is pervasive in mathematics, it is less common in the pricing literature. We refer the reader to Corless et al. [44] for a thorough review of its properties and provide only a brief summary below.

Recall, the general (multi-valued) Lambert- W function $W(x)$, is defined as a solution to

$$W(x)e^{W(x)} = x.$$

When $x \in [-1/e, 0)$, this equation has two distinct real solutions. The branch $W_{-1}(\cdot)$ gives the solution that lies in $(-\infty, -1]$. The other branch $W_0(\cdot)$ gives the solution in $[-1, \infty)$, but will not be needed in our work. Both branches are illustrated in the left panel of Fig. 2.2.

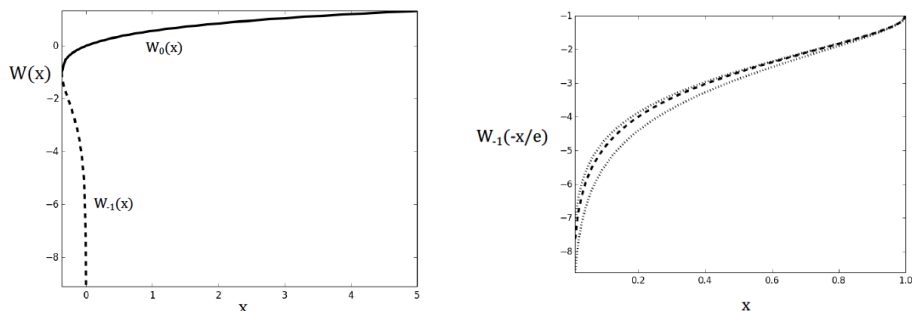


Figure 2.2: The left panel shows the two real branches of the Lambert- W function, $W_0(\cdot)$ (solid black), and $W_{-1}(\cdot)$ (dashed). Our bounds depend upon the $W_{-1}(\cdot)$ branch (rescaled), as shown in right panel, and which can be upper and lower bounded via Chatzigeorgiou [29] (dotted).

To build intuition, we encourage the reader to think of $W_{-1}(\cdot)$ as analogous to

the natural logarithm, $\log(\cdot)$. Indeed, like $W_{-1}(x)$, $\log(x)$ is defined as a solution to an equation, namely,

$$e^{\log(x)} = x.$$

For a handful of values, both $W_{-1}(\cdot)$ and $\log(\cdot)$ can be evaluated exactly. For example, $W_{-1}(-1/e) = -1$, $\log(1) = 0$, and $\lim_{x \rightarrow 0} W_{-1}(x) = \lim_{x \rightarrow 0} \log(x) = -\infty$. For most values, however, both functions must be evaluated numerically. Fortunately, evaluating an expression using $W_{-1}(\cdot)$ is numerically no more difficult than evaluating a similar expression using $\log(\cdot)$.

Moreover, the natural logarithm provides simple bounds on $W_{-1}(\cdot)$. Indeed, Chatzigeorgiou [29] proves that for $0 < x \leq 1$,

$$-1 - \sqrt{2 \log(1/x)} - \log(1/x) \leq W_{-1}\left(-\frac{x}{e}\right) \leq -1 - \sqrt{2 \log(1/x)} - \frac{2}{3} \log(1/x). \quad (2.1)$$

(Recall $W_{-1}(\cdot)$ is defined on $[-1/e, 0)$, so that this inequality spans its domain.) The right panel in Fig. 2.2 illustrates these bounds and shows they are quite tight.

2.3 The Value of Personalized Pricing over Single Pricing

In this section, we provide tight upper and lower bounds on the value of personalized pricing over single pricing using simple statistics and/or shape assumptions on F . We begin by upper bounding the value of personalized pricing using the *scale* (S), and *margin* (M), defined respectively as:

$$S := \frac{\inf\{k \mid F(k) = 1\}}{\mu}, \quad M := 1 - \frac{c}{\mu}.$$

These two statistics are unit-less and can be thought of as (rescaled) measurements of the maximal valuation and per unit cost. More specifically, S is the ratio of the largest valuation in the market to the average valuation. By construction, $S \geq 1$, and measures the maximal dispersion of valuations. By contrast, $M = \frac{\mu-c}{\mu} \in [0, 1]$, and can be interpreted as the margin of a unit sold at a price equal to the mean valuation.

Before stating our tight bound, we introduce a transformation that reduces the problem of bounding the value of personalization for a product with $c > 0$ and $\mu > 0$ to an equivalent problem with $c = 0$ and $\mu = 1$. This reduction is used repeatedly throughout the paper.

Lemma 2.3.1 (Reduction to Zero Costs and Unit Mean). *Let $V \sim F$, and let the distribution of $\frac{1}{\mu-c}(V - c)$ be denoted by F_c . Then,*

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} = \frac{\mathcal{R}_{PP}(F_c, 0)}{\mathcal{R}_{SP}(F_c, 0)}.$$

Moreover, if the scale and margin of F are S and M , respectively, then the mean, scale, and margin of F_c are $\mu_c = 1$, $S_c = \frac{S+M-1}{M}$, and $M_c = 1$, respectively.

The key to the following bound is that $\mathcal{R}_{SP}(F, 0)$ directly yields a bound on the tail behavior of F . Indeed, for any price $p > 0$, $p\bar{F}(p) \leq \mathcal{R}_{SP}(F, 0)$ by definition, and thus $\bar{F}(p) \leq \mathcal{R}_{SP}(F, 0)/p$. We use this result repeatedly in what follows, terming it the *pricing inequality*:

$$\bar{F}(x) \leq \frac{\mathcal{R}_{SP}(F, 0)}{x}, \quad \forall x > 0. \quad (\text{Pricing Inequality})$$

This inequality drives the following lemma.

Lemma 2.3.2 (Bounding $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$ using the Scale and Margin). *For any F with scale S*

and margin M , we have

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq -W_{-1} \left(\frac{-M}{e(S + M - 1)} \right).$$

Moreover, this bound is tight.

Proof. First, suppose $c = 0$ and $\mu = 1$. Then, $\mathcal{R}_{PP} = 1$ and $M = 1$. Since $\mu = 1$, $\bar{F}(S) = 0$, i.e., $0 \leq V \leq S$, a.s. Using the tail integral formula for expectation, we have that

$$\mathcal{R}_{PP} = \int_0^S \bar{F}(x) dx \tag{2.2}$$

$$\leq \mathcal{R}_{SP} + \int_{\mathcal{R}_{SP}}^S \bar{F}(x) dx \quad (0 \leq \mathcal{R}_{SP} \leq S) \tag{2.3}$$

$$\leq \mathcal{R}_{SP} + \int_{\mathcal{R}_{SP}}^S \frac{\mathcal{R}_{SP}}{x} dx \quad (\text{Pricing Inequality}) \tag{2.4}$$

$$= \mathcal{R}_{SP} + \mathcal{R}_{SP} \log \left(S \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \right) \quad (\text{since } \mathcal{R}_{PP} = 1).$$

Rearranging this inequality yields

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq 1 + \log \left(S \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \right). \tag{2.5}$$

We next use properties of $W_{-1}(\cdot)$ to simplify Eq. (2.5). Exponentiating both sides yields,

$$e^{\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}} \leq eS \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \iff \frac{1}{eS} \leq \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} e^{-\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}} \iff \frac{-1}{eS} \geq -\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} e^{-\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}} \tag{2.6}$$

Since $\frac{-1}{eS} \in [-1/e, 0)$ and the function $W_{-1}(\cdot)$ is non-increasing on this range, applying it to both sides of (2.6) and multiplying by -1 yields

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq -W_{-1} \left(\frac{-1}{eS} \right), \tag{2.7}$$

which proves the bound when $c = 0$ and $\mu = 1$, since $M = 1$.

To prove tightness, it suffices to construct a nonnegative random variable $V \sim F$ with $\mu = 1$ and scale S , such that $\mathcal{R}_{SP}(F, 0) = \frac{-1}{W_{-1}(\frac{-1}{eS})}$. For convenience, define $\alpha = \frac{-1}{W_{-1}(\frac{-1}{eS})}$, and notice, by definition of $W_{-1}(\cdot)$,

$$-\frac{1}{Se} = -\frac{1}{\alpha}e^{-\frac{1}{\alpha}} \iff \frac{\alpha}{S} = e^{1-\frac{1}{\alpha}} \iff \log\left(\frac{\alpha}{S}\right) = 1 - \frac{1}{\alpha} \iff \frac{1}{\alpha} = 1 + \log\left(\frac{S}{\alpha}\right). \quad (2.8)$$

Next consider a random variable with cCDF

$$\bar{F}_S(x) = \begin{cases} 1 & \text{if } x \in (0, \alpha] \\ \frac{\alpha}{x} & \text{if } x \in (\alpha, S] \\ 0 & \text{otherwise.} \end{cases}$$

Observe that F_S has mean 1, since

$$\mu = \int_0^S \bar{F}_S(x) dx = \alpha + \alpha \log\left(\frac{S}{\alpha}\right) = \alpha \left(1 + \log\left(\frac{S}{\alpha}\right)\right) = 1, \quad (2.9)$$

by Eq. (2.8). By inspection, F_S has scale S . Finally, for any $x \in (\alpha, S]$, $x\bar{F}_S(x) = \alpha$, and for any other x , $x\bar{F}_S(x) \leq \alpha$. Hence, $\mathcal{R}_{SP}(F, 0) = \alpha$, and, thus, the bound is tight for F_S .

For a general $c > 0$ and $\mu \neq 1$, use Lemma 2.3.1 to reduce to the case that $c = 0$, $\mu_c = 1$, $M_c = 1$, and $S_c = \frac{S+M-1}{M}$. Lemma 2.3.1 and Eq. (2.7) then imply that $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} = \frac{\mathcal{R}_{PP}(F_c, 0)}{\mathcal{R}_{SP}(F_c, 0)} \leq -W_{-1}\left(\frac{-1}{eS_c}\right)$. Replacing S_c proves the upper bound. Create a tight distribution by scaling F_{S_c} (defined above) by $\mu - c$ and shifting by c . \square

The described tight distribution is a truncated Pareto distribution on $[\alpha, S]$ for some $\alpha \in [c, S]$, which satisfies $\bar{F}_S(x) \propto 1/x$ on its support (see left panel Fig. 2.3). In the auction literature, this distribution is sometimes called the “equal-revenue”

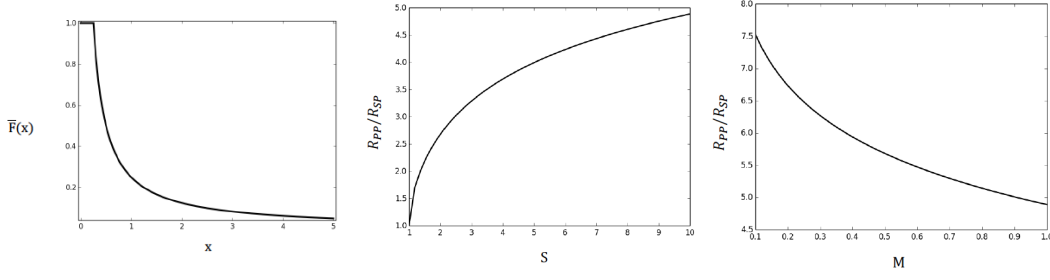


Figure 2.3: The left panel shows the tight distribution of Lemma 2.3.2 when $M = 1$ and $S = 5$. The middle panel shows the bound in Lemma 2.3.2 when $M = 1$ and as S varies between 1 and 10. The right panel shows the bound in Lemma 2.3.2 when $S = 5$ and as M varies between 0.1 and 1.0.

distribution, since all prices in $[\alpha, S]$ yield the same single-pricing profit. Thus, one optimal pricing strategy for this distribution is to price at $p = \alpha$ and sell to *all* customers.

In the middle and right panels of Figure 2.3, we plot the bound of Lemma 2.3.2 versus M and S . Intuitively, as the scale increases, valuations become more dispersed and personalization offers greater potential value, as seen in the middle panel. On the other hand, increasing the margin with a fixed mean is equivalent to decreasing the cost per unit. As discussed above, an optimal single-pricing strategy has the same market share as idealized personalized pricing under the tight distribution. Thus, in the right panel, as margin increases, the profits of both idealized personalized pricing and single pricing increase at the same rate, and their relative ratio decreases. We stress that this behavior crucially depends on the properties of the tight distribution.

Remark 2.3.1. *Many of our subsequent proofs utilize techniques similar to the proof of Lemma 2.3.2. Consequently, we highlight some of its high-level features before proceeding. First, the proof is centered around an integral representation of a moment of V (in this case μ) in terms of the cCDF \bar{F} (cf. Eq. (2.2)). The key step is to point-wise upper bound $\bar{F}(x)$ at each x . For $x \leq \mathcal{R}_{SP}$, the tightest bound possible is simply 1 (cf. Eq. (2.3)). For $x \geq \mathcal{R}_{SP}$, we use the Pricing Inequality (cf. Eq. (2.4)). The tight distribution is constructed by constructing a valid cCDF \bar{F} that simultaneously*

makes each of these point-wise bounds tight. The remaining steps are simple algebraic manipulation. Thus, the three key elements are an integral representation in terms of the cCDF, point-wise bounds on the cCDF, and identifying a single distribution which simultaneously matches all point-wise bounds. \square

Bounds Incorporating the Coefficient of Deviation

A drawback of Lemma 2.3.2 is that the bound becomes vacuous as the scale $S \rightarrow \infty$. The issue is that S , alone, cannot distinguish between markets where most customers have relatively similar valuations (which may be relatively low or high) and markets where customer valuations vary widely. We next provide more descriptive upper bounds on the value of personalized pricing by incorporating a measure of the market's heterogeneity, i.e., the typical dispersion in valuations. Specifically, we define the *coefficient of deviation* of F by

$$D := \frac{E[|V - \mu|]}{2\mu}.$$

By construction, $D \in [0, 1]$ since $E[|V - \mu|] \leq E[|V|] + \mu = 2\mu$ by the triangle inequality. Intuitively, D is the (rescaled) mean absolute deviation of V . Mean absolute deviation (or MAD) is a common measure of a random variable's dispersion, similar to standard deviation. Intuitively, when D is small, we expect most valuations to be close to μ , and, hence, the value of personalization to be small. By contrast, when D is large, we expect there to be larger dispersion in valuations, and, hence, the potential value of personalization to be much larger.

This intuition is not entirely correct as we shall see below. In fact, when D is *very* large and S is finite, there is a boundary effect; F is approximately a two-point distribution concentrated near c and μS , and single-pricing strategies are very effective. A single price can be used to capture the high valuation customers, while

the low valuation customers are simply ignored since their potential profitability is near zero. Consequently, for very large D , the value of personalization is, in fact, low.

This qualitative description is formalized in Theorem 2.3.1 which upper bounds the value of personalization in terms of S , M , and D . The theorem partitions the space of markets into three distinct regimes depending on the magnitude of D and provides distinct bounds for each regime. Specifically, we define the three regimes by

(L) Low Heterogeneity: $0 \leq D \leq \delta_L$

(M) Medium Heterogeneity: $\delta_L \leq D \leq \delta_M$

(H) High Heterogeneity: $\delta_M \leq D \leq \delta_H$,

where $\delta_L, \delta_M, \delta_H$ are constants that depend on M and S :

$$\delta_L := -\frac{M \log\left(\frac{S+M-1}{M}\right)}{W_{-1}\left(\frac{-1}{e^{\frac{S+M-1}{M}}}\right)}, \quad \delta_M := \frac{M \log\left(\frac{S+M-1}{M}\right)}{1 + \log\left(\frac{S+M-1}{M}\right)}, \quad \delta_H := \frac{M(S-1)}{S+M-1}.$$

The following lemma proves these regimes form a true partition:

Lemma 2.3.3 (Partitioning the Range of D). *Given F with scale S and margin M , the coefficient of deviation of F satisfies $0 \leq D \leq \delta_H$. Moreover, $0 \leq \delta_L \leq \delta_M \leq \delta_H$.*

Equipped with Lemma 2.3.3, we can state Theorem 2.3.1, the main upper bound of this section.

Theorem 2.3.1 (Bounding $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$ using D). *For any F with scale S , margin M , and coefficient of deviation D , we have the following:*

a) *If $0 \leq D \leq \delta_L$, then*

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{-W_{-1}\left(\frac{\frac{D}{M}-1}{e}\right)}{1 - \frac{D}{M}}. \quad (\text{Low Heterogeneity})$$

b) If $\delta_L \leq D \leq \delta_M$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{M \log \left(\frac{S+M-1}{M} \right)}{D}. \quad (\text{Medium Heterogeneity})$$

c) If $\delta_M \leq D \leq \delta_H$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq -W_{-1} \left(\frac{-1}{e^{\left(\frac{S+M-1}{M} \right) \left(1 - \frac{D}{M} \right)}} \right). \quad (\text{High Heterogeneity})$$

Moreover, for any S, M, D there exists a valuation distribution F with scale S , margin M and coefficient of deviation D such that the corresponding bound is tight.

Theorem 2.3.1 gives a complete, closed-form upper bound on the value of personalized pricing for any distribution in terms of its scale, margin, and coefficient of deviation. The bound is defined piecewise, but is continuous (cf. Fig. 2.4). Note that the bound captures the intuition that the value of personalization increases as D increases for small to moderate D , but also captures the boundary behavior as D becomes very large. The maximal point in Fig. 2.4, at the transition between the low and medium regimes, corresponds exactly to the bound in Theorem 2.3.2. When S is infinite, $\delta_L = 1$ and Theorem 2.3.1 reduces to simply Theorem 2.3.1(a). The bound is neither convex nor concave as a function of D .

We also observe that our bound in Figure 2.4 can be significantly above or below e , the uniform bound proven for *monotone hazard rate* (MHR) distributions in Barlow et al. [14] and Hartline et al. [69]. Further, although the value of personalized pricing can be infinite, our refined analysis characterizes precisely when classes of distributions lead to a low values of personalized pricing. Finally, the bound can easily be further upper-bounded using the approximations in Eq. (2.1) to avoid the Lambert- W function, but at the cost of tightness. The approximate bound is:

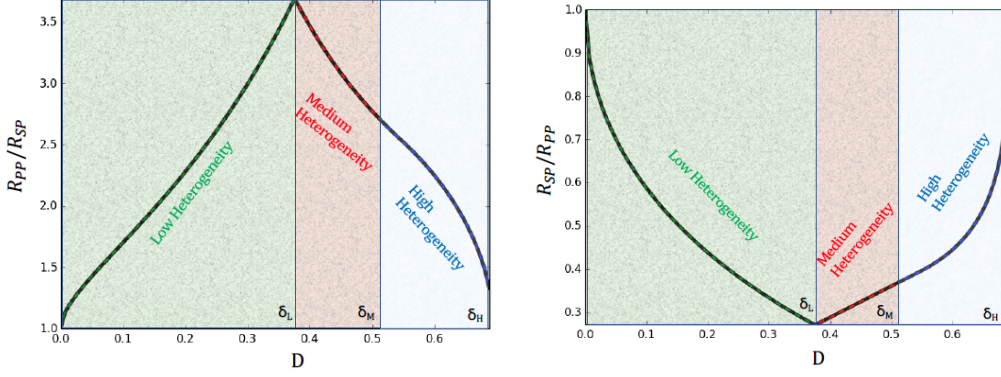


Figure 2.4: The left panel plots the bound from Theorem 2.3.1 as a function of D with $S = 4$ and $M = .9$. The right panel plots the inverse of this bound, which we note is convex.

a) If $0 \leq D \leq \delta_L$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{1 + \sqrt{2 \log \left(\frac{1}{1 - \frac{D}{M}} \right) + \log \left(\frac{1}{1 - \frac{D}{M}} \right)}}{1 - \frac{D}{M}}.$$

b) If $\delta_L \leq D \leq \delta_M$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{M \log \left(\frac{S+M-1}{M} \right)}{D}.$$

c) If $\delta_M \leq D \leq \delta_H$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq 1 + \sqrt{2 \log \left(\frac{S+M-1}{M} \left(1 - \frac{D}{M} \right) \right) + \log \left(\frac{S+M-1}{M} \left(1 - \frac{D}{M} \right) \right)}.$$

Single-Pricing Guarantee: An alternative interpretation of Theorem 2.3.1 is that the reciprocal of the bound is a tight guarantee on the performance of single-pricing relative to idealized personalized-pricing. In other words, the single-pricing strategy is guaranteed to earn at least the given percentage of the idealized personalized pricing profits. This perspective, i.e., interpreting single-pricing as an approximation to idealized personalized pricing, is common in the approximation algorithm literature.

We plot this guarantee, i.e., the reciprocal of the bound in Theorem 2.3.1, in the right panel of Fig. 2.4. Perhaps surprisingly, the reciprocal appears convex as a

function of D . We prove this formally in Lemma 2.3.4 and leverage this observation later in Section 2.4.

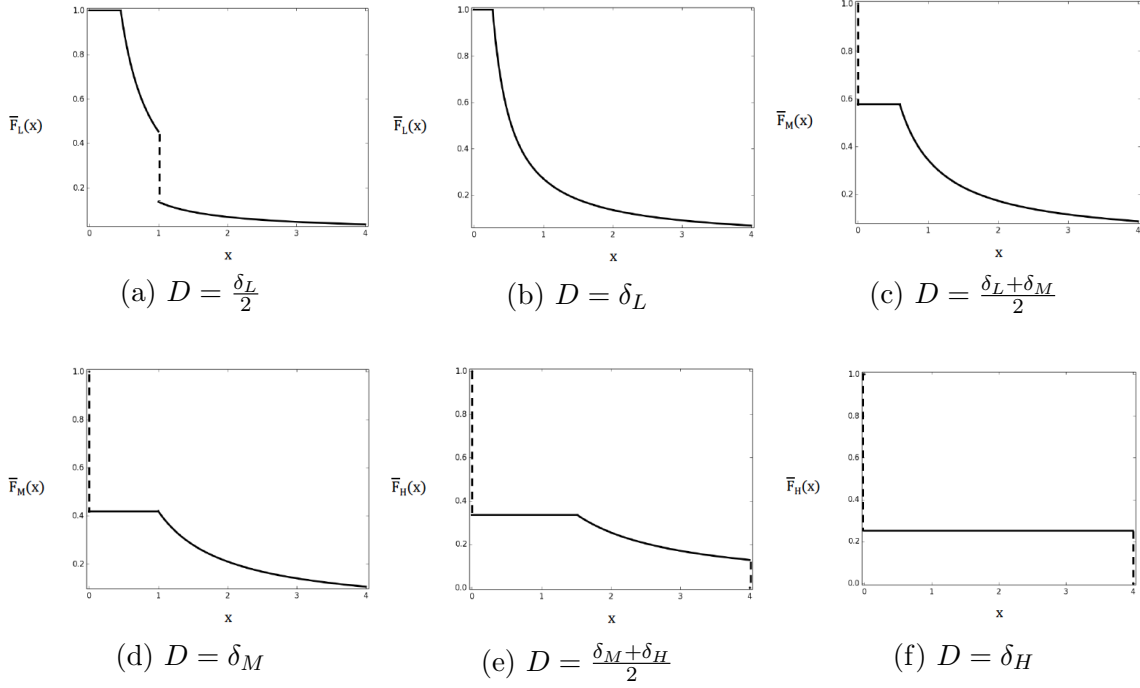
Lemma 2.3.4 (Convexity of the Single-Pricing Guarantee). *For any S , M , and D , let $\alpha(S, M, D)$ denote the reciprocal of the bound on the value of personalized pricing in Theorem 2.3.1. Then $\alpha(S, M, D)$ is a convex function in D .*

Tight Distributions: Like Lemma 2.3.2, Theorem 2.3.1 is a tight bound. The distribution which achieves the bound depends on the regime but is not unique. See Fig. 2.5 for typical examples and Lemma B.2.4 in the appendix for explicit formulas. In all three regimes, a worst-case distribution can be constructed from a mixture of a two-point distribution and truncated Pareto distributions; what differs between the regimes is the placement and sizes of these components. We show in the course of proving Theorem 2.3.1 that any price along the truncated Pareto section is an optimal price for the single-pricing strategy. These results generalize a folklore result from the auction literature that the Pareto distribution represents the worst-case valuation distribution (where S and D are unrestricted).

Although the forms of the tight distributions differ by regime and are not unique, it is instructive to consider a class of them as a function of D and, in particular, study how they evolve as D increases with all other parameters are fixed. We focus on $c = 0$ and $\mu = 1$ as in Fig. 2.5.

- When $D = 0$, the bound of Theorem 2.3.1 is 1 and the unique tight distribution is a point-mass on μ . For $D > 0$ but small, this point mass stretches into two Pareto curves above and below the mean. Every point along the Pareto curve below the mean is an optimal point at which to price, whereas no point along the Pareto curve above the mean is optimal (cf. Fig. 2.5a).
- As D grows towards δ_L , the Pareto curve above the mean rises to meet the curve below the mean. They join when $D = \delta_L$ as illustrated in Fig. 2.5b.

Figure 2.5: Tight distributions for Theorem 2.3.1 in each regime when $S = 4$, $\mu = 1$ and $M = 1$.



Every point along the resulting single curve is an optimal price, and we recover the tight distribution of Lemma 2.3.2. This point is the transition between the low and medium regimes, and yields the greatest value of personalized pricing.

- As D grows past δ_L , boundary effects force mass to begin to pool at zero, and the single Pareto curve begins to shrink with the left most end point tending away from 0 and back towards μ . Again, every point along the Pareto curve is an optimal point at which to price (cf. Fig. 2.5c).
- When $D = \delta_M$, all mass below μ is contained in a point mass on 0. The Pareto curve extends from μ to μS , and pricing at any point along it is optimal (cf. Fig. 2.5d).
- As D grows past δ_M , boundary effects intensify and force mass to pool on both zero and μS . Past μ is an increasingly short, flat Pareto curve, along which every point is optimal (cf. Fig. 2.5e).
- Finally, the distribution converges to a two-point distribution on 0 and μv_{max}

as illustrated in Fig. 2.5f.

Asymptotics Finally, from a theoretical point of view, one might seek to characterize the value of personalized pricing as D approaches its extreme values $D \rightarrow 0$ or $D \rightarrow \delta_H$. In particular, we will see in Section 2.4 that the first limit also provides insight into the performance of certain third-degree price discrimination tactics. These limits are below:

Corollary 2.3.1 (Asymptotic Behavior). *For any S, M, D , let $1/\alpha(D, M, S)$ denote the bound from Theorem 2.3.1. Then,*

a) As $D \rightarrow 0$,

$$\frac{1}{\alpha(S, M, D)} = 1 + \sqrt{2\frac{D}{M}} + O\left(\frac{D}{M}\right).$$

b) As $D \rightarrow \delta_H$,

$$\frac{1}{\alpha(S, M, D)} = 1 + \sqrt{2\frac{S+M-1}{M}} \cdot \sqrt{\delta_H - \frac{D}{M}} + O\left(\delta_H - \frac{D}{M}\right).$$

In both cases, $\frac{1}{\alpha(S, M, D)}$ approaches its limit like the square root of the difference from the boundary.

Lower Bounds on the Value of Personalized Pricing

In this subsection, we complement our upper bounds on the value of personalized pricing with closed form lower bounds. Such lower bounds are helpful in identifying when personalized pricing techniques are *necessary* to achieve strong revenue guarantees. Unfortunately when only S, M , and D are given, no non-trivial lower bound can be derived on the value of personalized pricing over single pricing. It can easily be seen that there exists a two point distribution with one point on zero that obtains any arbitrary, fixed S, M , and D , but for which the value of personalized pricing

over single pricing is 1. To avoid these pathological two point distributions, we will require to impose two additional assumptions about the distributions *shape*, namely that the distribution is *unimodal* and *left-skew*.

Definition 2.3.1. *A distribution F is unimodal if there exists some α such that F is a convex function on $(-\infty, \alpha]$ and concave function on (α, ∞) .*

Definition 2.3.2. *A unimodal distribution is left-skew if its mode α precedes its mean μ i.e. if $\alpha \leq \mu$.*

We note that the class of left-skew distributions subsumes the more commonly studied class of symmetric distributions. Further, many natural distributions including uniform, normal, exponential and others are unimodal and left-skew. We will lean on these two shape assumptions to give lower bounds on the value of personalized pricing over single pricing. To the best of our knowledge, this bound is the first of its kind, yielding generic separation between the revenue of a single price strategy over a general class of distributions.

Theorem 2.3.2 (Lower Bounding $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$). *For any unimodal, left-skew distribution F with margin M , and coefficient of deviation D ,*

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \geq \begin{cases} \frac{1}{1 - \frac{D}{M}} & \text{if } \frac{D}{M} \leq 1/3, \\ \frac{8 \frac{D}{M}}{(1 + \frac{D}{M})^2} & \text{if } \frac{D}{M} \geq 1/3 \end{cases}.$$

Remark 2.3.2. *Theorem 2.3.2 follows by leveraging the tail convexity of the cCDF of unimodal distributions via the following geometric fact: the area of any rectangle inscribed in a right triangle is no more than half the total area of the triangle. For example, consider Fig. 2.6(a) which gives a visual representation of \mathcal{R}_{SP} (area of shaded rectangle) in relation to \mathcal{R}_{PP} (total area under the curve). By unimodality, the cCDF is convex on $[\alpha, \infty)$. If (SP) uses a price p^* which is greater than the mode*

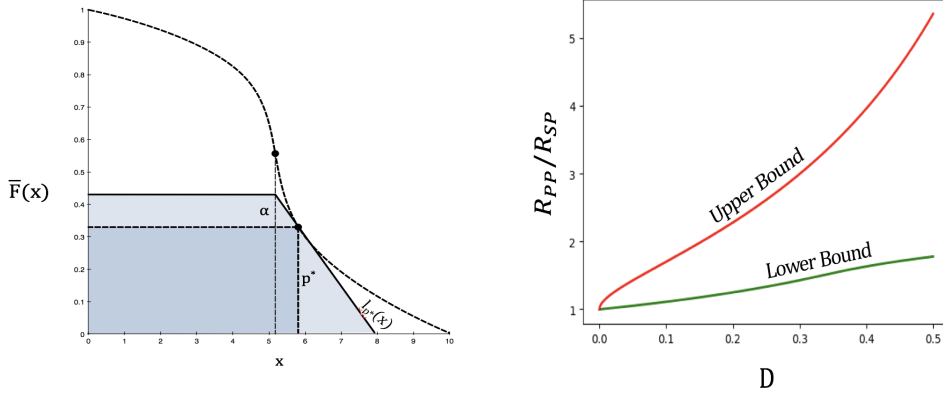


Figure 2.6: The left panel shows the revenue of a single price strategy (dark colored square) in relation to the supporting lower bound (light colored trapezoid) for a unimodal cCDF. The right panel plots the guarantee of Theorem 2.3.2 (green) against the upper bound of Theorem 2.3.1 (red).

α , the revenue a single price strategy earns from customers with valuations higher than α is exactly the area of the center rectangle in Fig. 2.6(a), which is inscribed in the right triangle made by the supporting line of the convex curve of \bar{F} , and our simple geometric fact applies. Finally, the additional assumption of left-skewness enforces that a non-trivial amount of the mass exists above the mode where separation between \mathcal{R}_{SP} (area of the rectangle) and \mathcal{R}_{SP} (area under the curve) applies.

We note that Theorem 2.3.2 is tight as function of a technical parameter, $\lambda := \frac{\int_0^\alpha \bar{F}(x) dx}{\mu}$, and tightness is achieved by an appropriately shifted uniform distribution. This tightness in the parameter lambda translates to tightness in the input parameter D when the valuations are symmetric via an application Lemma B.2.1. Further, Theorem 2.3.2 exhibits the correct dependencies as $D \rightarrow 0$, where the bound tends to 1 as expected. When $D \rightarrow 1$, the bound tends to 2 which is best possible for the class of symmetric distributions and corresponds to the case when the cCDF is fully convex. For comparison, in Fig. 2.6(b) we plot the lower bound in Theorem 2.3.2 against the upper bound Theorem 2.3.1(a) when $D \in [0, .5]$.

2.4 From Third-Degree to First-Degree Price Discrimination

As mentioned in the introduction, idealized personalized pricing is a strategy that hinges on two assumptions: 1) the firm can charge potentially distinct prices to every customer and 2) the firm is omniscient. In this section, we analyze how each of these assumptions contributes to the value of personalized pricing. In particular, we compute the value of personalized pricing over k -market segmentation and feature-based pricing. We then use both of these results to bound the value of personalized pricing over feature-based market segmentation. Our bounds yield insight into how these strategies “converge” to idealized personalized pricing as $k \rightarrow \infty$ or predictive accuracy increases. Said another way, they quantify both the value of the operational capability to charge a continuum of prices and the value of additional predictive accuracy.

Feature-Based Pricing

In this section, we study the value of personalized pricing over *feature-based pricing*. From a practical point of view, feature-based pricing approximates a host of third-degree price discrimination strategies in common use. For example, student discounts are a form of feature-based pricing where \mathbf{X} is a binary indicating that the customer is a student. More generally, in online retailing settings, sellers often have access to rich contextual information for each customer from her cookies, such as demographics, browsing history, etc., that can be used to personalize the offered price via a custom coupon.

Clearly, if one can perfectly predict V from \mathbf{X} , feature-based pricing is equivalent to idealized personalized pricing. Typically, however, \mathbf{X} is not rich enough to predict V perfectly, entailing some loss in profits. Thus, from a theoretical point of view,

$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{XP}}$ quantifies the benefits of additional *information*, i.e., the benefit of observing a richer set of features that enable perfect prediction. We will be most interested in the rate at which $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{XP}} \rightarrow 1$ as the information in \mathbf{X} increases. Loosely speaking, this rate describes the predictive accuracy needed from a model to guarantee a given percentage of idealized personalized pricing profits.

Formally, we assume that the seller has trained a prediction model using historical data such that for any realization of \mathbf{X} , the seller knows the conditional distribution $V \mid \mathbf{X} \sim F_{V|\mathbf{X}}$. Let $\mu(\mathbf{X}) \equiv E[V \mid \mathbf{X}]$, and define the residual ϵ of the model by $V = \mu(\mathbf{X}) + \epsilon$. Note, by construction, $E[\epsilon \mid \mathbf{X}] = 0$ almost surely.

As an example, suppose the valuations follow the well-known logit model, i.e., a customer's valuation is a linear combination of that customer's features, the offered price and an idiosyncratic error with a logistic distribution. For this model, the conditional distribution is known precisely, and

$$\mathbb{P}(V - c \geq p \mid \mathbf{X}) = \frac{1}{1 + e^{-(\beta_0 p + \boldsymbol{\beta}^\top \mathbf{X})}}.$$

Our assumption is that the seller has learned the coefficients $\beta_0, \boldsymbol{\beta}$.

A first, perhaps obvious, observation is that given \mathbf{X} , it is not optimal to price at $E[V \mid \mathbf{X}]$. To the contrary, one should price at the optimal price for the conditional distribution $F_{V|\mathbf{X}}$. This essentially proves Lemma 2.4.1.

Lemma 2.4.1 (Relating Feature-Based Pricing and Single-Pricing). *For any joint distribution $F_{\mathbf{X}V}$, we have $\mathcal{R}_{XP}(F_{\mathbf{X}V}, c) = E[\mathcal{R}_{SP}(F_{V|\mathbf{X}}, c)]$.*

In Theorem 2.4.1, we use this observation in conjunction with our previous bounds on \mathcal{R}_{SP} to bound the value of idealized personalized pricing over feature-based pricing under mild assumptions on the form of the valuation distribution.

Theorem 2.4.1 (Idealized Personalized Pricing vs. Feature-Based Pricing). *Suppose that $V = E[V \mid \mathbf{X}] + \epsilon$ where the residual ϵ satisfies $E[|\epsilon| \mid \mathbf{X}] = E[|\epsilon|]$. Suppose*

further that there exists δ with $0 < \delta < 1$ such that $\mu(\mathbf{X}) \geq \frac{c}{1-\delta}$ almost surely. Then,

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{XP}(F_{\mathbf{X}V}, c)} \leq \frac{1}{\alpha\left(\bar{S}, M, \frac{E[|\epsilon|]}{2\mu}\right)},$$

where $\alpha(S, M, D)$ denotes the reciprocal of bound in Theorem 2.3.1 and $\bar{S} = \frac{M(S-1)+\delta}{\delta}$.

Interestingly, when the coefficient of deviation of V is in the ‘low heterogeneity’ regime, the bound in Theorem 2.4.1 has same form as the one in Theorem 2.3.1, except that the MAD of V is replaced by the MAD of the residual noise. (The scale does not appear in either bound.) This implies that the value of additional feature information in this regime can be directly measured by how much the residual MAD is reduced.

We consider the assumption in Theorem 2.4.1 that $E[|\epsilon| \mid \mathbf{X}] = E[|\epsilon|]$ to be quite mild. This assumption underlies many predictive models used in practice, including the logit model described above. Indeed, for the logit model, ϵ is a (centered) logistic random variable, *independent* of \mathbf{X} , so the above assumption holds directly. Similar remarks hold for other regression-based models with independent errors.

Similarly, we also consider assumption that $\mu(\mathbf{X}) \geq \frac{c}{1-\delta}$ almost surely to be mild. It holds, e.g, whenever $\bar{F}(c + \delta(\mu - c)) = 1$, i.e., the effective support of V is well-separated from c . We stress that when D is in the low-heterogeneity regime, $\alpha(S, M, D)$ does not depend on S . Thus, even if δ is very small (causing \bar{S} to be very large), the above bound is unaffected.

Intuitively, one can think of ϵ as the residual in the non-parametric regression $V = \mu(\mathbf{X}) + \epsilon$. If \mathbf{X} is very informative for V , we expect ϵ , and hence, $\frac{E[|\epsilon|]}{2\mu}$ to be small. As we acquire more features, $E[|\epsilon|] \rightarrow 0$ and valuations can be predicted more accurately from these features. Simultaneously, the value of personalized pricing over feature-based pricing tends to 1. In sum, Theorem 2.4.1 provides a simple formula

for benchmarking the quality of a predictive model for pricing and understanding the benefit of additional features.

We next provide numerical experiments to demonstrate that the bound in Theorem 2.4.1 is reasonably accurate in terms of magnitude and shape, as a function of the coefficient of deviation of the residual ϵ . Specifically, we generate valuations according the model

$$V = 10 + \sum_{i=1}^{10} X_i + \epsilon \quad (2.10)$$

where each $X_i \sim N(0, 1)$, and ϵ is either a (centered) Logistic, Gumbel, or (shifted) Exponential distribution with standard deviation of 1 and mean zero.

In our experiment, we suppose the seller only knows the first k features, and thus $\mu(X_1, \dots, X_k) = 10 + \sum_{i=1}^k X_i$. The corresponding error term, ϵ_k , has distribution of $X_{k+1} + \dots + X_{10} + \epsilon$. In Fig. 2.7, we plot the actual value of personalized pricing as a function of the number of features available to the seller and the bound from Theorem 2.4.1. Observe the bound is quite illustrative in both magnitude and shape.

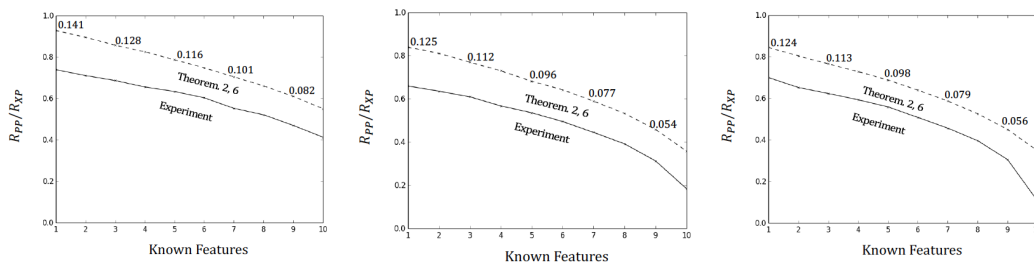


Figure 2.7: Illustrates the decreasing benefit of idealized personalized pricing over feature-based pricing as the number of incorporated model features increases. The numbers above the curve denote the scaled MAD of the unexplained noise, $\frac{E[|\epsilon|]}{2\mu}$, for every other k . We plot Logistic (left), Gumbel (middle), and Exponential (right) noise, respectively.

Finally, Theorem 2.4.1 can be used prescriptively to determine the accuracy of a predictive model needed to guarantee a given percentage of idealized personalized-pricing profits. In particular, if a monopolist seeks $(1 - \beta)$ -fraction of the idealized personalized pricing profits, it suffices to construct a predictive model with enough

features so that $\alpha(\bar{S}, M, \frac{E[|\epsilon|]}{2\mu}) \geq 1 - \beta$. By Corollary 2.3.1, for small β , this amounts to $\frac{E[|\epsilon|]}{2\mu} < \frac{\beta^2}{2(1-\beta)^2} = O(\beta^2)$. Although this analysis is based on an upper-bound, Fig. 2.7 suggests the general dependence on $E[|\epsilon|]$ is correct, i.e., to halve the gap (between \mathcal{R}_{PP} and \mathcal{R}_{XP}), one needs 4 times the predictive accuracy for small β .

Market Segmentation

In the section, we study the value of personalized pricing over *k-market segmentation*. From a practical point of view, *k-market segmentation* approximates settings in which the monopolist's ability to predict customer valuations is good, but her ability to charge different prices to different customers is limited. For instance, the monopolist may be constrained to only offer 10%, 20%, or 30%-off coupons (rather than a continuum of prices), but can identify the valuation of a customer accurately enough to place them in one of these buckets.

From a theoretical point of view, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}}$ quantifies the benefit of an *operational* capability – the ability to offer a continuum of prices rather than a finite set. Intuitively, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}} \rightarrow 1$ as $k \rightarrow \infty$. We will be most interested in the rate at which this convergence occurs. Intuitively, this rate characterizes how many segments one must use to guarantee a given percentage of idealized personalized pricing profits.

We first establish a simple lemma on the structure of the optimal segmentation.

Lemma 2.4.2 (Structure of Optimal Segmentation). *There exists an optimal segmentation s_0, \dots, s_{k+1} and pricing p_1, \dots, p_k for \mathcal{R}_{kP} such that $s_i = p_i$ for $i = 1, \dots, k$.*

Proof. If $s_i < p_i$, then increasing s_i to p_i does not affect revenue in segment $[s_i, s_{i+1})$ and can only increase revenue in segment $[s_{i-1}, s_i)$. \square

Using this simple observation, we can explicitly compute the value of personalized pricing over *k-market segmentation* for uniform random variables.

Lemma 2.4.3 (*k*-Market Segmentation and Uniform Valuations). *Let $V \sim F$ be a uniform random variable supported on $[0, t]$. Then, $\frac{\mathcal{R}_{PP}(F, 0)}{\mathcal{R}_{kP}(F, 0)} = 1 + \frac{1}{k}$.*

The first part of Theorem 2.4.2 below proves that in the worst-case, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}}$ is also $1 + \tilde{O}\left(\frac{1}{k}\right)$ which matches the uniform case (up to logarithmic factors). The second part shows that with a mild assumption on F , i.e., its support is well-separated from c , we can additionally drop these logarithmic factors. Thus, in light of Lemma 2.4.3 the worst-case rate of convergence of Theorem 2.4.2 is essentially tight.

Theorem 2.4.2 (Idealized Personalized Pricing vs. *k*-Market Segmentation). *For any valuation distribution F with scale S and margin M , and for any $k \in \mathbb{N}$,*

a) If F has coefficient of deviation D , then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kP}(F, c)} \leq 1 + \frac{1 + \log\left(\frac{S+M-1}{M} \left(1 + \frac{D}{M}(k+1)\right)\right)}{k} = 1 + \tilde{O}\left(\frac{1}{k}\right).$$

b) If there is some $\delta > 0$ such that $\overline{F}(c + \delta(\mu - c)) = 1$, then

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kP}(F, c)} \leq 1 + \frac{\log\left(\frac{S}{\delta}\right)}{k}.$$

The proof of Theorem 2.4.2 constructs a (suboptimal) segmentation strategy by geometrically partitioning the valuation space. In each segment, we then leverage our previous bounds on the value of personalized pricing over single-pricing to bound the profit.

Interestingly, the dependence $1 + O\left(\frac{1}{k}\right)$ appears typical for many distributions. In Figure 2.8, we plot the exact ratio $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}} - 1$ for three different distributions, as well as our bound from Theorem 2.4.2 with $c = 0$ and $\delta = \frac{1}{\mu}$. Specifically, the first panel considers shifted Beta distributions, i.e., $V \sim \text{Beta}(\alpha, 3) + 1$ for $\alpha = 0.1, 1.325, 2.55, 3.775, 5.0$. The second panel considers truncated exponential distributions, i.e., $V \sim \max(\min(\text{Exp}(\alpha), 2), 1)$ for $\alpha = 0.5, 1.0, 1.5, 2.0$. Finally, the third

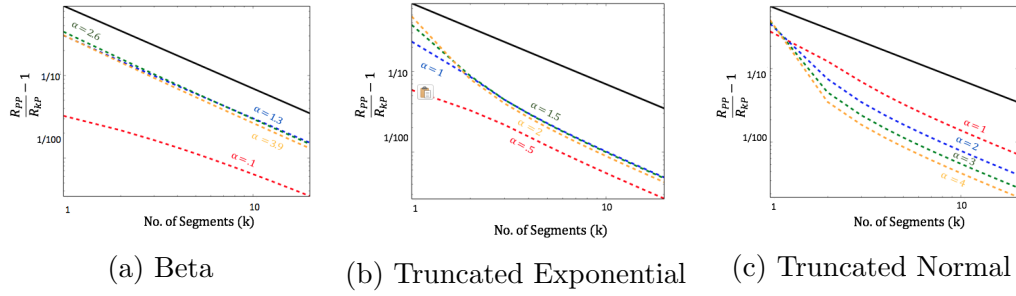


Figure 2.8: Illustrates the decreasing benefit of idealized personalized pricing over optimally segmenting into k groups for particular distributions (dotted lines), and our distribution-agnostic bound (solid line). Note the log-scales.

panel considers truncated normal distributions, i.e., $V \sim \max(\min(\text{Norm}(1, \alpha), 2), 1)$, for $\alpha = 1, 2, 3, 4$. Note the log-scales.

In each case, the dependence on k appears similar, and matches the dependence in Theorem 2.4.2. Intuitively, this behavior can be explained in the following way. As we segment into smaller pieces, any distribution with a continuous density appears locally uniform on each segment. Example 2.4.3 establishes that the convergence rate for a uniform matches Theorem 2.4.2 up to constant factors, suggesting that, at least for large k , the rate should also be approximately tight for many distributions.

Further, Theorem 2.4.2 can be used prescriptively to determine the number of segments necessary to guarantee a specific percentage of idealized personalized pricing profits. Namely, a monopolist seeking to guarantee $1 - \beta$ -fraction of the idealized personalized pricing profits needs to use $k \geq (1/\beta - 1) \log(\frac{S}{\delta}) = O(1/\beta)$ by part (b) of Theorem 2.4.2. While this upper bound does not provide a tight analysis, Fig. 2.8 suggests the dependence on k is approximately tight for small β , i.e, to halve the relative gap to idealized personalized pricing, one needs twice as many segments.

Finally, we note that when F is given and discretely supported on N points, the optimal k -market segmentation strategy can be efficiently computed via a simple dynamic programming algorithm. We defer the details to Appendix B.3 and shall return to this observation in Section 2.4.

Feature-Based Market Segmentation

In this section, we study the value of personalized pricing over *feature-based market segmentation*. Feature-based market segmentation closely resembles real-world data-driven personalization strategies where sellers are both constrained by the number of the prices they can offer, and must predict customer valuations from data. In this way, feature-based market segmentation synthesizes the two previous models of personalized pricing discussed in this section. Formally, feature-based market segmentation is equivalent to feature-based pricing with the restriction that the seller can offer only k distinct prices. As in the previous section, we assume the seller has learned the conditional distribution $F_{V|\mathbf{X}}$ from data, and we let $\mu(\mathbf{X}) = E[V | \mathbf{X}]$ and define the residual ϵ by $V = \mu(\mathbf{X}) + \epsilon$.

We bound the value of personalized pricing over feature-based market segmentation by separately considering the loss from limited price flexibility and the loss from the prediction error in the valuation model. We measure *loss* as the difference in profit between a personalized pricing strategy ((kP), (XP), or (kXP)) and idealized personalized pricing. The following theorem states that one can bound the loss of feature-based market segmentation by the loss of two more powerful strategies: feature-based pricing, and k -market segmentation on a noiseless market.

Theorem 2.4.3 (Idealized Personalized Pricing vs. Feature-Based Market Segmentation). *As above, let $V = \mu(\mathbf{X}) + \epsilon$, and suppose \mathbf{X} and ϵ are independent. Then*

$$\underbrace{\mathcal{R}_{PP}(F, c) - \mathcal{R}_{kXP}(F_{\mathbf{X}V}, c)}_{\text{Loss of } kXP} \leq \underbrace{\mathcal{R}_{PP}(F, c) - \mathcal{R}_{kP}(F_{\mu(\mathbf{X})}, c)}_{\text{Loss of } kP \text{ on noiseless market}} + \underbrace{\mathcal{R}_{PP}(F, c) - \mathcal{R}_{XP}(F_{\mathbf{X}V}, c)}_{\text{Loss of } XP}.$$

If $\mathcal{R}_{kP}(F_{\mu(\mathbf{X})}) + \mathcal{R}_{XP}(F_{\mathbf{X}V}, c) > \mathcal{R}_{PP}(F, c)$, this implies

$$\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kXP}(F_{\mathbf{X}V}, c)} \leq \frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kP}(F_{\mu(\mathbf{X})}, c) + \mathcal{R}_{XP}(F_{\mathbf{X}V}, c) - \mathcal{R}_{PP}(F, c)}.$$

Unlike Theorem 2.4.1, Theorem 2.4.3 requires that ϵ is independent of \mathbf{X} . This is a stronger assumption, but, as discussed previously, is satisfied by many common valuation models, such as the logit model.

We emphasize that Theorem 2.4.3 is tight when the firm can offer an infinite number of prices, or when the prediction error of the valuation model vanishes. To see this note that for any market F , when a firm has infinite price flexibility $\lim_{k \rightarrow \infty} \mathcal{R}_{kXP}(F_{\mathbf{X}V}, c) = \mathcal{R}_{XP}(F_{\mathbf{X}V}, c)$ and $\lim_{k \rightarrow \infty} \mathcal{R}_{kP}(F_{\mu(\mathbf{X})}) = \mathcal{R}_{PP}(F_{\mu(\mathbf{X})}, c) = \mathcal{R}_{PP}(F, c)$. When the firm's valuation model has perfect prediction accuracy, i.e. $\epsilon \sim 0$, $\mathcal{R}_{kXP}(F_{\mathbf{X}V}, c) = \mathcal{R}_{kP}(F_{\mu(\mathbf{X})})$ and $\mathcal{R}_{XP}(F_{\mathbf{X}V}, c) = \mathcal{R}_{PP}(F, c)$. In general Theorem 2.4.3 forms a theoretical framework through which a firm can analyze the performance of personalized pricing strategies. In particular, the bound neatly decouples the profit loss from limited price flexibility, through analyzing (kP) strategies on the noiseless market $F_{\mu(\mathbf{X})}$, and the prediction error through analyzing the (XP) strategy.

Operationally, a monopolist may use Theorem 2.4.3 to guide the tuning of personalized pricing strategies. Consider a firm that learns a valuation model $\mu(\cdot)$ from customer data. Suppose $\mu(\mathbf{X})$ is supported on n points (one can imagine that $\mu(\cdot)$ is trained on n sampled data points), then the optimal (kP) pricing strategy on the noiseless market $F_{\mu(\mathbf{X})}$ can be computed in time $O(kn^2)$ by dynamic programming, using the algorithm in Appendix B.3. Thus $\mathcal{R}_{kP}(F_{\mu(\mathbf{X})})$ can be computed exactly by a firm after training the model. Since valuations are not observable, the distribution of ϵ is not obtainable. Instead, we can use the basic statistics of ϵ (typically given as output after training the prediction model) and then bound \mathcal{R}_{XP} using Theorem 2.4.1 and Theorem 2.3.1. With an exact computation of $\mathcal{R}_{kP}(F_{\mu(\mathbf{X})})$ and a bound on \mathcal{R}_{XP} , the firm can apply Theorem 2.4.3 to study the performance of the feature-based market segmentation strategy. This allows the seller to reason about where exactly the inefficiency is arising, and decide whether to increase the number of prices/segments,

or invest in additional data gathering to reduce the prediction error, or both.

Finally, we note that the proof Theorem 2.4.3 is constructive, and implies a heuristic for setting feature-based market segmentation strategies: compute the optimal k -market segmentation $\{s_i\}_{i=1}^{k+1}$ for the noiseless market $F_{\mu(\mathbf{x})}$, use it to generate the segments \mathcal{X}_i , then perform price experimentation to learn the prices that maximize $p_i \Pr(s_i + \epsilon \geq p_i)$, and offer that price on each segment. While both the partition into segments $\{\mathcal{X}_i\}_{i=1}^k$, and the prices offered on each segment $\{p_i\}_{i=1}^k$ may be sub-optimal, such a strategy is guaranteed to earn more than $\mathcal{R}_{XP}(F_{\mathbf{x}V}, c) + \mathcal{R}_{kP}(F_{\mu(\mathbf{x})}, c) - \mathcal{R}_{PP}(F, c)$ by Theorem 2.4.3.

2.5 Extensions

In Section 2.3 we developed upper and lower bounds on the ratio $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)}$ that depended on the margin, scale, and coefficient of deviation and some shape assumptions. In this section we extend those bounds in two directions: (i) Idealized personalized pricing is, as the name suggests, unobtainable. For many practitioners feature-based pricing strategies is the more relevant model of personalized pricing. In these cases, deciding whether or not to invest the time and resources necessary to implement sophisticated feature-based pricing strategies boils down to understanding the possible revenue lift over setting a single price i.e., $\frac{\mathcal{R}_{XP}(F, c)}{\mathcal{R}_{SP}(F, c)}$, not $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)}$. To address these concerns we prove a novel extension theorem that transforms bounds on $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)}$ into tighter bounds on $\frac{\mathcal{R}_{XP}(F, c)}{\mathcal{R}_{SP}(F, c)}$. (ii) We provided bounds on the value of personalized pricing over single-pricing in terms of the valuation distribution's coefficient of deviation. Although the coefficient of deviation enjoys properties that make it amenable to closed-form analysis, in principle, any statistic might be used. In the second half of this section, we show how to compute bounds on the value of personalized pricing over single-pricing for other natural statistics such as the variance, geometric mean,

and quantiles via infinite dimensional linear programming duality.

Extending Bounds on $\frac{\mathcal{R}_{PP}(F,c)}{\mathcal{R}_{SP}(F,c)}$ to $\frac{\mathcal{R}_{XP}(F,c)}{\mathcal{R}_{SP}(F,c)}$

In this subsection we turn our attention to the ratio between feature-based pricing and a single price strategy. Following the notation in Section 2.4, we will show how to transform the upper and lower bounds from Section 2.3 on the value of personalized pricing into bounds $\frac{\mathcal{R}_{XP}(F,c)}{\mathcal{R}_{SP}(F,c)}$. Intuitively, the performance of a feature-based pricing strategy will depend on how close the trained model $\mu(\mathbf{X})$ is to the true conditional valuations, $V|\mathbf{X}$. When the prediction is relatively noiseless, we expect $\mathcal{R}_{XP} \approx \mathcal{R}_{PP}$ and the bounds in Theorems 2.3.1 and 2.3.2 can be applied without modification. On the other hand, when the model error ϵ is impactful, the relative power of feature-based pricing compared to single price strategies should be diminished. Theorem 2.5.1 formalizes this intuition under the same mild assumptions discussed in Sections 2.4 and 2.4.

Theorem 2.5.1 (Feature-Based Pricing vs. Single Pricing). *Suppose $V = \mu(\mathbf{X}) + \epsilon$, $\mu(\mathbf{X})$ is supported on $[c, \infty)$, ϵ is unimodal, left-skew, supported on $[-c, \infty)$ and $E[\epsilon] = 0$. Further suppose $\mu(\mathbf{X})$ and ϵ are independent. Let $V_{\mu+\epsilon} = E[\mu(\mathbf{X})] + \epsilon$ and $M_{\mu+\epsilon}$, $D_{\mu+\epsilon}$ be the margin and coefficient of deviation of $V_{\mu+\epsilon}$, respectively.*

$$\frac{1 - \frac{D_{\mu+\epsilon}}{M_{\mu+\epsilon}}}{-W_{-1}\left(\frac{\frac{D_{\mu+\epsilon}}{M_{\mu+\epsilon}} - 1}{e}\right)} \frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{\mathcal{R}_{XP}(F, c)}{\mathcal{R}_{SP}(F, c)} \leq \frac{2 - \frac{D_{\mu+\epsilon}}{M_{\mu+\epsilon}}}{2} \frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{SP}(F, c)}$$

For ease of exposition suppose $M_{\mu+\epsilon} = 1$, and observe that both $\frac{1 - D_{\mu+\epsilon}}{-W_{-1}\left(\frac{D_{\mu+\epsilon} - 1}{e}\right)}$ and $\frac{2 - D_{\mu+\epsilon}}{2}$ equal 1 when $D_{\mu+\epsilon} = 0$, and are less than 1 when $D_{\mu+\epsilon} > 0$. This matches our intuition, when $D_{\mu+\epsilon} = 0$, $\mathcal{R}_{XP}(F, c) = \mathcal{R}_{PP}(F, c)$. As the coefficient of deviation of the error increases, the relative power of feature-based pricing decreases resulting in tighter upper bounds and weaker lower bounds on the ratio between feature-based

pricing and a single price strategy. Next, note that Theorem 2.5.1 allows for *generic* transformations between bounds on the value of personalized pricing and $\frac{\mathcal{R}_{XP}(F,c)}{\mathcal{R}_{SP}(F,c)}$, any upper or lower bound on $\frac{\mathcal{R}_{PP}(F,c)}{\mathcal{R}_{SP}(F,c)}$ can be plugged into Theorem 2.5.1. In the next two subsection we show to compute generic bounds on the value of personalized pricing.

Upper Bounds Based upon General Moments

In Section 2.3 we derived tight, closed form upper bounds on the value of personalized pricing over single-pricing in terms of the valuation distribution's coefficient of deviation, in this subsection we will show how to compute bounds for other moments. Specifically, we seek to upper bound the value of personalization in terms of the scale, mean, and a specific moment $E[f(V)]$ of F , where $f(\cdot)$ is a known fixed function. For example, when $f(v) = \frac{|v-\mu|}{2\mu}$, this moment is equal to the coefficient of deviation D of F . When $f(v) = \frac{(v-\mu)^2}{\mu^2}$, this moment equals the squared coefficient of variation of F . Finally, when $f(v) = \mathbb{I}(v \geq \hat{p})$, this moment equals the fraction of the market that purchases at price \hat{p} , e.g., an incumbent price, under F . By possibly redefining f , i.e., shifting by a constant, we can with out loss of generality assume that $E[f(V)] = 0$.

The key idea of our approach is to formulate a continuous mathematical optimization program that explicitly computes the value of personalized pricing over distributions which satisfy the above constraints. To build intuition, we first consider

the case when $c = 0$ and $\mu = 1$. Consider the optimization problem

$$\begin{aligned}
& \sup_{y, d\mathbb{P}_v} 1/y & (2.11) \\
& \text{s.t.} \quad \int_0^S d\mathbb{P}_v = 1, \quad d\mathbb{P}_v \geq 0, \quad \forall v \in [0, S] \\
& \quad \int_0^S v d\mathbb{P}_v = 1 \\
& \quad \int_0^S f(v) d\mathbb{P}_v = 0 \\
& \quad y \geq p \int_0^S \mathbb{I}(v \geq p) d\mathbb{P}_v \geq 0, \quad \forall p \in [0, S].
\end{aligned}$$

The optimization variables above are \mathbb{P}_v , which represents the measure of V , and y , which represents the single-pricing profit. The first constraint ensures that \mathbb{P}_v is a valid probability measure. The second constraint ensures the mean of the distribution is 1. The third constraint ensures that $E[f(V)] = 0$. Finally, the last family of (infinite) constraints ensures that y is at least the revenue achieved by pricing at p for any $p \in [0, S]$. At optimality, y will equal the optimal single price revenue by choice of objective, and \mathbb{P}_v will be the distribution with smallest possible single price revenue. Therefore, (2.11) computes a tight upper bound on the value of personalized pricing.

Unfortunately, with both infinite constraints and infinite variables, problem (2.11) appears computationally challenging. Theorem 2.5.2 below provides an alternate mathematical program with a finite number of variables and infinite number of constraints that provides an upper bound on the value of personalized pricing (and the solution value of (2.11)). We present the theorem in the case of general $c > 0$ and $\mu > 0$ for completeness.

Theorem 2.5.2 (General Bound on Value of Personalized Pricing). *Let F be any distribution with scale S , margin M and mean μ that satisfies $E[f(v)] = 0$ for a fixed, known $f(\cdot)$. Let $0 = p_0 < p_1 < \dots < p_{N-1} < p_N = \frac{S+M-1}{M}$ be a discretization of the interval $[0, \frac{S+M-1}{M}]$*

and define

$$\begin{aligned}
z^* &:= \max_{\theta, \lambda, \mathbf{Q}} \theta + \lambda_1 & (2.12) \\
s.t. \quad & \sum_{j=0}^N Q_j = 1, \quad Q_j \geq 0, & j = 0, \dots, N, \\
& \theta + \lambda_1 v + \lambda_2 f(v\mu M + \mu(1 - M)) \leq \sum_{j=0}^{k-1} p_j Q_j, & \forall v \in [p_{k-1}, p_k], \quad k = 1, \dots, N \\
& \theta + \lambda_1 \frac{S + M - 1}{M} + \lambda_2 f(S\mu) \leq \sum_{j=0}^N p_j Q_j.
\end{aligned}$$

Then, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq 1/z^*$.

The proof of Theorem 2.5.2 involves three steps: (i) rewriting (2.11) as a minimization over z , (ii) applying continuous linear optimization duality, and (iii) discretizing the resulting dual program. We defer the details to Section B.2.

Unlike our previous bounds, Eq. (2.12) depends on the mean μ . For moment functions $f(\cdot)$ that are scaled relative to μ , however, this dependence often disappears. For example, in the case of coefficient of deviation where $f(V) = \frac{|V-\mu|}{2\mu} - D$, $f(v\mu M + \mu(1 - M)) = \frac{|M(v-1)|}{2} - D$ which does not depend on μ .

The tractability of Eq. (2.12) depends crucially on the function $f(\cdot)$. We argue that despite the infinite number of constraints, this problem can be solved efficiently, both theoretically and practically, as long as one can efficiently identify an optimizer of

$$\max_{v \in [p_{k-1}, p_k]} \lambda_1 v + \lambda_2 f(v\mu M + \mu(1 - M)) \quad (2.13)$$

for every k and every λ_1, λ_2 . Indeed, if one can identify such an optimizer, it is possible to separate over these constraints efficiently.

Namely, given a candidate solution $(\theta, \lambda_1, \lambda_2)$, solve Eq. (2.13) for each k and let v_k^* denote the optimizer. Then check if

$$\theta + \lambda_1 v_k^* + \lambda_2 f(v_k^* \mu M + \mu(1 - M)) \leq \sum_{j=0}^{k-1} p_j Q_j, \quad k = 1, \dots, N. \quad (2.14)$$

If all these constraints are satisfied, $(\theta, \lambda_1, \lambda_2)$ is feasible for the original set of infinite constraints. Otherwise, if the k^{th} constraint is violated, it defines a separating hyperplane that separates $(\theta, \lambda_1, \lambda_2)$ from the feasible region.

Using standard machinery, the above separation routine can be used in conjunction with the ellipsoid method to prove polynomial time tractability of Eq. (2.12) whenever finding an optimizer to Eq. (2.13) is polynomial time. Alternatively, this separation routine also yields a constraint generation procedure that can be combined with the dual simplex method for a practically efficient, but not necessarily polynomial time, algorithm. Specifically, we sequentially add violated constraints by checking Eq. (2.14) and resolving. If after some iteration no constraints are violated, we terminate. Otherwise, we repeat. If desired, we can terminate the algorithm early by computing the maximum constraint violation s in Eq. (2.14) for the current solution, and observing that $(\theta - s, \lambda_1, \lambda_2)$ is feasible in Eq. (2.12). Thus, $1/(\theta + \lambda_1 - s)$ is a valid upper bound on $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$. We employ this constraint generation procedure with early termination in Section B.1 and Section B.1 in the appendix.

In summary, the tractability of Eq. (2.12) hinges on the ability to optimize Eq. (2.13). We highlight three important cases where an optimizer to Eq. (2.13) can be found efficiently:

- **$f(\cdot)$ is convex** In this case, whenever $\lambda_2 \geq 0$, then the objective of Eq. (2.13) is convex, and the optimizer is one of the end points p_{k-1} or p_k . If $\lambda_2 < 0$, then Eq. (2.13) is a univariate, concave maximization problem which can be solved with standard techniques.
- **$f(\cdot)$ is concave** This case is similar to the above case.
- **$f(\cdot)$ is piecewise linear** When $f(\cdot)$ is a piecewise linear function with a known, small number of pieces, the optimizer of Eq. (2.13) occurs either at one of these knots or at an endpoint of the interval $[p_{k-1}, p_k]$.

We stress Eq. (2.12) is thus tractable whenever $f(\cdot)$ has one of these forms.

Finally, we discuss choosing the number of discretization points N in Eq. (2.12). Notice for any $N \geq 1$, Theorem 2.5.2 provides a valid upper bound on the price of personalization. By contrast, an alternate approach might be to discretize Eq. (2.11) directly, i.e., restrict at-

tention to measures \mathbb{P}_v supported on p_0, \dots, p_N . While the optimal value of this discretized problem does *not* yield a valid bound on personalized pricing, it does provide a lower bound on the optimal value to Eq. (2.11), and, thus, bounds the potential value of increasing N . Consequently, a heuristic approach to choosing N might be to increase N until the relative gap between $\frac{1}{z^*}$ and the optimal value of this discretized problem is sufficiently small, say, less than 1%. We apply this approach in Fig. B.1 to study the sensitivity of our approach to the choice of N .

Lower Bounds Based upon General Moments

In Section 2.3 we derived tight, closed form lower bounds on the value of personalized pricing over single-pricing in terms of the valuation distribution's coefficient of deviation under a shape assumption, namely that the distribution was symmetric and left-skew. As in Theorem 2.3.1 with upper bounds, Theorem 2.3.2 depended on particular properties of the coefficient of deviation that make it amenable to closed-form analysis, however in principle, any statistic might be used. In this subsection, we describe an efficient procedure to generate lower bounds on the value of personalized pricing over single-pricing in terms of more general moments, under the same shape assumptions (namely unimodality and left-skewness) mirroring the approach in Section 2.5.

Specifically, our approach will be to derive upper bounds on the revenue of a single pricing strategy in terms of the mean, a specific moment $E[f(V)]$ of F where $f(\cdot)$ is a known fixed function, over the class of unimodal, left-skew distributions. We will follow the approach of Popescu [88] who show how to solve general moment bound problems over the class of unimodal distributions using semi-definite programming. The only new idea will be to show how to appropriately discretize the space of possible prices so as yield bounds efficiently.

Suppose $\mu = 1$, $c = 0$, we are given an upper bound S , and fix a price $p \in [0, S]$. Then the revenue of a single price strategy using price p , over the class of distributions which are unimodal, left-skew, and satisfy $E[f(X)] = q$, is upper bounded by the following

optimization problem:

$$\begin{aligned}
(P_p) \quad \Delta_p = \sup_{d\mathbb{P}_v} \quad & p \int_p^S d\mathbb{P}_v \\
\text{s.t.} \quad & \int_0^S v d\mathbb{P}_v = 1 \\
& \int_0^S f(v) d\mathbb{P}_v = q \\
& \int_0^S d\mathbb{P}_v = 1, \quad d\mathbb{P}_v \geq 0, \quad \forall v \in [0, S], \quad \mathbb{P} \text{ unimodal, left-skew}
\end{aligned} \tag{2.15}$$

The optimization variables above are \mathbb{P}_v , which represents the measure of V . The first constraint ensures the mean of the distribution is 1. The second constraint ensures that the moment condition $E[f(V)] = q$ is met. Finally, the last family of (infinite) constraints ensures that $d\mathbb{P}_v$ is a valid distribution in the class of unimodal, left-skew distributions. In order to obtain valid upper bounds we may solve $O(\frac{S}{\delta^2})$ problems for prices $p_i = (1 + \delta)^i - 1$. Then defining $\Delta := \max_i \Delta_{p_i}$, the following lemma will show $(1 + \delta)\Delta$ is a valid upper bound on $\mathcal{R}_{SP}(F)$ for any F in the class.

Lemma 2.5.1. *Let \mathcal{F} be some class of distributions which is contained in the class superset of all distributions supported on $[0, S]$ with mean 1. Fix $\delta < 1$ and let $\Delta = \{(1 + \delta)^i - 1 | (1 + \delta)^i \leq S + 1\}$. Then,*

$$\sup_{F \in \mathcal{F}} \mathcal{R}_{SP}(F, 0) \leq (1 + \delta) \sup_{F \in \mathcal{F}} \max_{p \in \Delta} \mathcal{R}_{SP}(F, 0, p)$$

Thus obtain valid lower bounds on $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}}$ it suffices to solve a number of programs (P_p) for fixed p . Unfortunately, we do not know how to directly solve (P_p) . Instead, we apply the general framework of Popescu [88] (namely the results in Section 4.3, Lemma 4.3) to obtain a dual representation of (P_p) that can be solved in polynomial time via semi-definite programming techniques as long as f satisfies a *piece-wise polynomial* (pp) condition. Specifically the results of Popescu [88] hold when the both the objective and the moment constraints are pp, however one can easily observe that for the objective in (P_p) , $E[p1\{X \geq p\}]$ is

piece-wise constant. For completeness we state the dual problem of (P_p) :

$$\begin{aligned}
(D_p) \quad z_p = \min_{\lambda_1, \lambda_2} \quad & \lambda_1 + \lambda_2 q \\
\text{s.t.} \quad & \int_1^t x \lambda_1 + f(x) \lambda_2 - p \mathbf{1}\{x \geq p\} dx \geq 0 \quad \forall t \in [1, S] \\
& t \lambda_1 + f(t) \lambda_2 - p \mathbf{1}\{t \geq p\} \geq 0 \quad \forall t \in [0, 1]
\end{aligned} \tag{2.16}$$

All together, for any desired accuracy parameter $\delta > 0$, we may obtain an near optimal upper bound on the revenue of \mathcal{R}_{SP} which implies near optimal lower bounds on the value of personalized pricing over single pricing. Following Popescu [88], we note that can be solved exactly as a semi-definite program (SDP). However, as SDPs can be quite complicated and are not commonplace in all solvers, we believe it is useful to show how to solve as a linear program for a special subclass of moment functions $f(\cdot)$. In particular, if f is either convex or concave we may efficiently separate over the second set of constraints in by solving:

$$\min_{t \in [0, 1]} t \lambda_1 + f(t) \lambda_2 - p \mathbf{1}\{t \geq p\} \tag{2.17}$$

Further, since the first set of constraints is the integration of Eq. (2.17), there are most two critical points in $[1, S]$ where the function could be minimized, call them c_1, c_2 . Taking these points along with the boundary points reduces checking feasibility of a candidate solution over the first infinite set of inequalities to checking at most four constraints. From there we can apply standard machinery Theorem 2.5.2 to get a solution. The following proposition summarizes the above discussion.

Proposition 2.5.1 (General Lower Bounds on Value of Personalized Pricing). *Let F be any unimodal, left-skew distribution, distribution with scale S , and mean 1 that satisfies $E[f(v)] = 0$ for a fixed, known convex/concave function $f(\cdot)$. Let c_1, c_2 be solutions to $t \lambda_1 + f(t) \lambda_2 - p \mathbf{1}\{t \geq p\} = 0$ and c_3 be a solution to $\frac{d}{dt} t \lambda_1 + f(t) \lambda_2 - p \mathbf{1}\{t \geq p\} = 0$. Fix $\delta < 1$ and let $\Delta = \{(1 + \delta)^i - 1 | (1 + \delta)^i \leq S\}$ be a discretization of the interval $[0, S]$ and*

define

$$\begin{aligned}
(D_p) \quad z_p &= \min_{\lambda_1, \lambda_2} \lambda_1 + \lambda_2 q & (2.18) \\
s.t. \quad & \int_1^t x \lambda_1 + f(x) \lambda_2 - p 1\{x \geq p\} dx \geq 0 \quad \forall t \in \{1, c_1^{\mathbb{I}(1 \leq c_1 \leq S)}, c_2^{\mathbb{I}(1 \leq c_2 \leq S)}, S\} \\
& t \lambda_1 + f(t) \lambda_2 - p 1\{t \geq p\} \geq 0 \quad \forall t \in \{0, c_3^{\mathbb{I}(0 \leq c_3 \leq 1)}, 1\}
\end{aligned}$$

$$Then, \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \geq \frac{1}{(1+\delta) \max_{p \in \Delta} z_p}.$$

2.6 Conclusions

Increasingly rich consumer profiles and choice models enable retailers to personalize to consumers at finer and finer levels. However, building such tools comes at an investment cost in the form of technology, data scientists, marketing, etc. Motivated by this trade-off, we provide a framework to quantify the benefits of personalized pricing in terms of the features of the underlying market. In particular, we exactly characterized the value of personalized pricing over posting a single price for all customers in terms of the scale, coefficient of deviation, and margin of the valuation distribution in closed-form.

Using our closed-form bound, we are also able to bound the value of personalized pricing over certain third-degree price discrimination tactics that more closely mirror current practice. Specifically, we first provide an order optimal bound on the value of personalized pricing over k -market segmentation. Intuitively, this bound quantifies the benefit of the operational ability to set a continuum of prices rather than k fixed prices. We then provide a bound on the value of personalized pricing over feature-based pricing strategies. Intuitively, this second bound quantifies the benefit of obtaining additional market information or improving one's predictive model. Finally we leveraged these two bounds to study the performance of feature based mar-

ket segmentation, a strategy which closely models popular, data-driven personalized pricing strategies.

Overall, we believe that our results provide a rigorous foundation for analyzing pricing strategies in the context of personalization. Our results can be used both by researchers attempting to design algorithms for personalized pricing, as well as by managers seeking to implement or improve their pricing strategies.

Chapter 3

Pricing and Design of Loot Boxes

In this chapter we study the use of loot boxes in online video games. The contents of this section are joint work with *Adam Elmachtoub, Xiao Lei, and Ningyuan Chen*.

Online video games represent a multi-billion dollar industry, where more than half of revenue is from “micro-transactions” - small exchanges of real life currency for virtual items to be used in the game. One popular way to sell virtual items are via *loot boxes*, which are random bundles of virtual items, where the contents are revealed after purchase. In this work we consider how to design and price loot boxes, and compare their revenue against standard bundle selling and separate selling strategies. We show that when the number of items is large, carefully designed loot box strategies can asymptotically garner as much revenue as grand bundle selling while inheriting many nice properties of separate selling.

In particular, in online environments where the game client has full information of the players current collection, we show unique loot box strategies that always allocate unique items to the player are asymptotically revenue optimal. In contrast, we show that loot box strategies that allocate items uniformly at random earn only $\frac{1}{e}$ of the optimal revenue asymptotically. We then extend these results to the case where there are multiple classes of items, where both prices and allocation probability need to be specified. We also show that when there are salvage costs, loot box selling dominates separate selling strategy. In both cases we obtain closed form expressions for the asymptotic revenues. Finally, we numerically validate our results for moderately sized catalogs of items, and propose heuristic policy with good performance.

3.1 Introduction

With the recent explosion of mobile gaming over the last decade, the video game industry standard has shifted towards the freemium model, where access to the game is freely given to customers and in-game virtual items can be acquired via micro-transactions. In other words, purchases of virtual items that help players in the game are made in exchange for small amounts of real currency (Forbes [58]). In many of these games, micro-transactions are conducted via a randomized mechanism known as a *loot box*. A loot box is a randomly filled bundle of virtual items, the contents of which are revealed after purchase. See Fig. 3.1 for an example. In online games such as *Dota 2*, *PlayerUnknown's Battlegrounds*, *Brawlhalla* and many others, loot box sales are a core source of revenue. In these games players purchase loot boxes which contain a random subset of virtual items such as character costumes, cosmetic upgrades, virtual cards, etc. In 2018 alone, more than \$30 billion dollars in sales were conducted via loot boxes (JuniperResearch [76]).



Figure 3.1: Depicted is a loot box offered by the popular video game *Dota 2*, which is called "treasure" in the game. The customer may purchase the loot box for \$2.49, after which they will receive one of the five items depicted on the screen.

More generally, the online gaming market is a particular rich area worthy of academic study. When selling virtual items for in-game use, the seller (typically the designer for the game) does not merely control the supply of items, but also conceives

of their existence and their value in the game. The seller has complete information and control in several interesting ways. Since the items are virtual, once the item is developed there is zero marginal cost for producing more units. Since the items only have value within the context of a particular game, the seller is allowed to fully control the resale market for the items. Since the customer interacts with the seller to play the game, the seller is fully aware of each customers' current inventory of items. Furthermore, sellers are essentially unrestricted in their choice of selling strategy. In light of these freedoms, it is then natural to ask: why does the gaming industry overwhelmingly choose to employ loot boxes to generate sales?

It is also important to note that, while popular and lucrative, loot box systems have been the subject of some recent controversy (Fingas [56]), where loot boxes have been characterized as a form of gambling and been subjected to regulatory investigation (Tassi [100]). In light of this negative publicity, in this work we study loot boxes as a revenue management tool and set out to understand *why* loot box selling has been so popular and successful. Specifically, we wish to understand why sellers would choose to implement loot box mechanisms over simple and effective selling mechanisms like *separate selling*, where every virtual item is sold at a single fixed price, or *grand bundle selling* where customers pay a fixed amount up front for all in-games items.

Towards this end, we propose the first rigorous revenue management framework for studying loot box strategies. We suppose that customers are utility maximizing, and characterize the revenue of two natural loot box strategies, as well as the underlying loot box pricing policies that achieve them. The two loot box policies we study are (i) *unique box* strategies, where the loot boxes are designed to always allocate an item the customer does not yet own, and (ii) *traditional box* strategies, where items are allocated uniformly at random regardless of whether or not the customer already owns a copy. As the number of items grows, we show, asymptotically, that

unique box strategies are fully revenue extracting and traditional box strategies are sub-optimal. Further, our asymptotically optimal loot box pricing policies are quite robust, depending only on the mean and variance of the valuation distribution. We then extend our results to capture many important and practical settings.

Our contribution can be described as follows:

1. We propose a natural model for loot box selling, which includes characterizing a rational customer's optimal purchasing policy. When the number of items is small, we show that there is no dominance relations between unique box, traditional box, grand bundling and separate selling strategies.
2. To overcome the general incomparability in the finite item setting, we consider the case where the number of items is large, an assumption which is often satisfied in the mobile gaming industry. We show that the revenue of an optimal unique box selling strategy asymptotically matches the revenue of an optimal grand bundle strategy (which is best-possible). On the other hand, the revenue of the popular traditional box strategies achieves only $\frac{1}{e}$ of the optimal revenue, and may earn even less than separate selling.
3. We then extend our loot box model to capture practical extensions including multiple item classes, loot boxes containing multiple items, and loot box strategies incorporating salvage systems. For loot boxes with multiple item classes, we exactly characterize the optimal loot box allocation policy, and show this policy has a natural form. We then show that previous revenue guarantees continue to hold when loot boxes are allowed to contain multiple items. Finally, when salvage systems are introduced, we show that both loot box strategies dominate separate selling, and get a complete order of the selling strategies asymptotically.
4. We conduct a set of numerical experiment which confirm the efficacy of loot box selling, even outside the asymptotic regime of our theoretical results. In

particular, for moderate sized catalogs of items and various valuation distributions, our loot box strategies achieve almost as much revenue as grand bundle selling, even when implementing heuristic prices based on our analysis.

Literature Review

As the gaming industry shifts towards mobile markets (Forbes [58]), loot box selling has come under increased scrutiny from the media, industry players, and regulators (Forbes [57], Apple [5], Tassi [100]), with some law makers asserting that loot box selling is a form of gambling (BusinessInsider [25]). In spite of this negative publicity, loot box selling is as popular and profitable as ever (GamesIndustry [61]). While media coverage has been extensive, there is comparatively little academic literature explaining why loot box selling has been so profitable. An emerging stream of literature in psychology has attempted give behavioral explanations for the effectiveness of loot box selling by connecting loot boxes with the larger literature on gambling (Drummond and Sauer [48], Zendle and Cairns [107]). We instead take operations approach and initiate the first rigorous mathematical treatment of loot boxes. While loot box selling has not, to the best of our knowledge, directly appeared in the revenue management literature, our work draws from and fits into several areas across operations management, computer science, and economics.

In the operations literature our work connects with the dual streams of papers on opaque selling and bundle selling. Opaque selling is the practice of selling goods where some feature of the product is hidden from the customer until after purchase, in loot box selling the allocation of the loot box is whats “opaque”. While traditionally opaque selling has been studied an inventory management tool, recent works Jerath et al. [74], Elmachoub and Wei [52], Elmachoub and Hamilton [51] have focused on opaque selling as tool to increase revenue by mitigating competition, and utilizing customer heterogeneity. Our work departs from these in two key ways, first in our

model the customer interacts with the mechanism sequentially, and the second there is no notion of customer between items. Our work also resembles and references the work on bundling. We compare our loot box selling mechanisms explicitly with the grand bundle mechanisms studied in the seminal work of Bakos and Brynjolfsson [10], who show that pure bundling extracts almost all of the consumer surplus. Mixed bundle strategies have been considered in recent work of Abdallah [1], Abdallah et al. [2]. In a sense, a loot box can be thought of as bundle of possible goods, of which only a single item is allocated. Mechanisms of this form are considered in Briest and Roglin [22] who study so called ‘unit demand bundles’, but in a static model, and with the aim of providing computational hardness results.

Next, there is a large and robust literature in the algorithmic game theory community which focuses on simple, approximately optimal selling strategies. In this vein, Hart and Nisan [67], Hart and Reny [68] study separate selling and grand bundle selling in a general static model, and show such strategies can approximate the optimal deterministic mechanism for buyers with additive valuations. In follow up work, Babaioff et al. [9] show that the better of these two strategies is a 6 approximation of the optimal deterministic mechanism. In our work we instead compare loot box selling against these simple mechanisms, and in a related sequential model. Further, Briest et al. [23] show how to set optimal prices for randomized mechanisms, which are menus of lotteries over the items, and prove the revenue of such mechanisms can greatly exceed the revenue of optimal deterministic mechanisms. Building on this work, Hart and Nisan [67], Briest et al. [24] show that lottery pricing can earn infinitely more revenue than any deterministic mechanism when the number of items is finite. Loot boxes are an inherently a randomized mechanism, and can be thought of a restricted lottery over subsets of the item, but the focus of our work is different. We are interested in comparing loot box selling against other practically prevalent mechanisms, not approximating the optimal mechanism in some class.

Finally, our work connects with the literature on *dynamic* mechanism design and sequential selling. In dynamic mechanism design, the agent interacts repeatedly with the seller over multiple periods, where typically some information about the state of the system is changing over time. In this literature, the focus tends to be strategic interaction between the buyer and seller, see for instance Pai and Vohra [85] or the excellent survey of Bergemann and Valimaki [16]. In this sense, loot box selling is a dynamic mechanism with trivial strategic interaction. In the computer science Chawla et al. [32, 31] study so called sequential posted price mechanisms, where an multiple agents arrive one at the time to the mechanism and are offered a price for a good. Along similar lines, in the operations literature Ferreira and Goh [55] study the dynamic assortment problem with strategic customer interaction. Our work sits between static and sequential mechanism design, where a single agent repeated interacts with the mechanism but all features of the mechanism are fixed over the course of the period.

3.2 Model and Preliminaries

We consider a profit maximizing monopolist selling N distinct, non-perishable, virtual items. A random customer's valuation for the items are described by a set of non-negative random variables $\{V_i\}_{i=1}^N$, where V_i are drawn i.i.d from F . The mean and variance of V_i are denoted by μ and σ^2 , respectively. Each customer has knowledge of all items in the seller's catalog, and every customer privately knows their own valuations V_i for $i = 1, \dots, N$. No customer desires more than one unit each item, meaning a customer's valuation for a second unit of item i is 0.

We now describe the sequence of events in our model. In each period t , let $S_t \subset [N]$ denote the set of unique items that the customer owns. The seller offers a loot box with price and allocation rule specified at period 0, and then at each period, customer

decides whether or not to purchases based on their valuations of items in $[N] \setminus S_t$. We assume customers are utility maximizing and will purchase if their expected utility is non-negative, otherwise the customer will not purchase and permanently leave the system. We discuss customer behavior more in details in Section 3.2.

We now highlight and justify some key assumptions in the model.

1. *Valuations for the items are i.i.d..*

The assumption of i.i.d. valuations is reasonable when the items are cosmetic and do not affect the balance of the game i.e. character skins/customization's, or when items are of similar impact i.e. cards of the same rarity, which is the case in many games that deploy boxes. The uniform retailing price in Figure 3.3b verifies this assumption to some extent. That said, there are situations when items are natural heterogeneous and can be separated into multiple classes based on rarity/game impact. We extend the model to address this case in Section 3.4.

2. *Goods are allocated by the loot box uniformly at random, and the allocation probabilities are known to the customers.*

The assumption of uniform allocation probabilities for loot boxes is natural for i.i.d. valued items, and is currently in practice for many loot box applications, i.e., Figure 3.2a. Still, one can imagine a seller manipulating the allocation probabilities for different items. We address this case in Section 3.4 and show that a revenue maximizing sellers optimal strategy is always to announce uniform allocation probabilities. The assumption that allocation probabilities are announced by the sellers is commonly satisfied in practice. Often sellers are forced to announce the allocation probabilities, either by government issued customer protection regulations [100], or by edict of the games distributor [5].

3. *A customer's valuation for a duplicate item is zero.*

The assumption that customer gain zero utility from duplicates of an item is

reasonable in the context of virtual items. For example in the case of duplicate cosmetic items, perhaps two of the same character skin, a second item offers no advantage over first. In some applications the seller includes a salvage mechanism i.e. a mechanism through which the customer can obtain value from duplicate items by trading them in for (possibly in game) currency. We discuss this case in Section 3.4.

4. *The seller is a monopolist.*

The seller, which is often the game designer, has the full control to the items in the game, including the quality and value. Also, items in a video game has zero value outside the game. Hence, it is reasonable to assume that the seller is a monopolist.

We consider four selling strategies in this framework: two forms of loot box selling, grand bundle selling, and separate selling.

1) Unique Box (UB): In the *unique box* strategy, the monopolist offers a loot box for a static price p , with the guarantee that after each purchase yields a new item that the customer does not yet own. The probability of receiving an item is 0 if $i \in S_t$, and $\frac{1}{|[N] \setminus S_t|}$ for $i \in [N] \setminus S_t$, i.e., uniform over all the items not currently owned by the customer. Fig. 3.2a shows an example of a unique box in a real game. We let $\mathcal{R}_{UB}(p)$ be the normalized revenue of a unique box strategy that uses price p , i.e.,

$$\mathcal{R}_{UB}(p) := \frac{p \times \mathbb{E}[\# \text{ Unique Loot Box Purchases}]}{N}$$

and let $\mathcal{R}_{UB} = \max_p \mathcal{R}_{UB}(p)$.

2) Traditional Box (TB): In the *traditional box* strategy, the monopolist offers a loot box for a static price p with the guarantee that the item allocated by the loot box is chosen with replacement, uniformly at random from the set $[N]$. This allows for the possibility that a customer receives duplicate items. Fig. 3.2b shows an example

of a traditional box in a real game. We let $\mathcal{R}_{TB}(p)$ be the normalized revenue of a unique box strategy that uses price p , i.e.

$$\mathcal{R}_{TB}(p) := \frac{p \times \mathbb{E}[\# \text{ Traditional Loot Box Purchases}]}{N}$$

and let $\mathcal{R}_{TB} = \max_p \mathcal{R}_{TB}(p)$.



(a) Unique Box



(b) Traditional Box

Figure 3.2: The left panel shows an implementation of a loot box in the popular online game *Dota 2*. The red square highlights that it is a unique box as the loot box always allocates a unique item. The right panel shows an implementation of a loot box in the online game *PlayerUnknown's Battlegrounds*. It allows duplicates and is a traditional box.

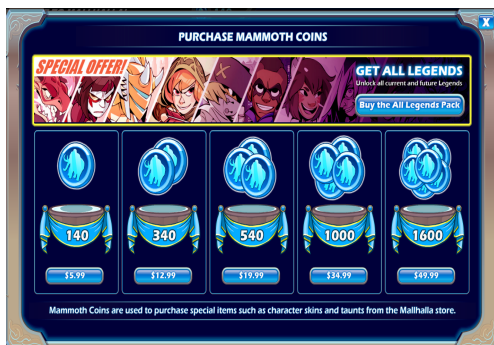
It is not clear which loot box strategy is better at first glance. Intuitively, customers may have higher valuation to a unique box, which induces a high retailing price. On the other hand, customer may purchase more than one traditional box to get a new item, which induces high selling volume. We shall compare and contrast these loot box models against two classic selling models: grand bundle selling and separate selling.

3) Grand Bundle (GB): In the *grand bundle* strategy, the monopolist offers a single bundle containing *all* N items for a static price Np . The normalized revenue of an optimal grand bundle strategy is then

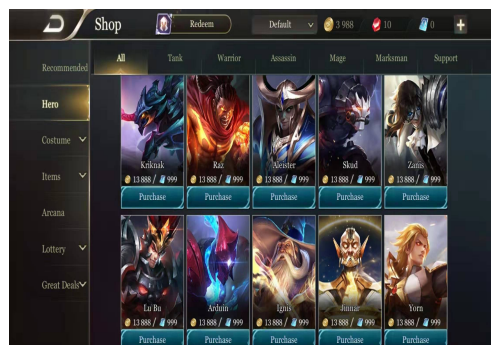
$$\mathcal{R}_{GB} := \max_p \frac{\Pr(\sum_i V_i \geq Np)}{N}.$$

Fig. 3.3a shows an example of grand bundle in practice.

4) Separate Selling (SS): In the *separate selling* strategy, the monopolist sells all items individually at the same price p . Since we assume the valuation V_i are i.i.d, the normalized revenue of an optimal separate selling strategy is then $\mathcal{R}_{SS} = \max_p p \Pr(V \geq p)$. Fig. 3.3b shows an example of separate selling with uniform prices.



(a) Grand Bundle



(b) Separate Selling

Figure 3.3: The left panel shows an implementation of grand bundle selling in the online game *Brawlhalla*. All items may be unlocked for a one time payment of \$19.99 via the *All Legends Pack* (although customers may also buy “Mammoth coins” and subsets of the items). The right panel shows an implementation of separate selling in the online game *Arena of Valor*. In this game each item (character) can be individually unlocked for a single payment.

While there are many more possible strategies in which to sell virtual items, we restrict our attention to these four as they capture the spirit of almost all strategies observed in practice. In Section 3.4, we confide a variety of extensions such as budget constraint, multi-class item valuation, loot boxes that allocate multiple items, and loot box selling with salvage system. Further, when the budget is infinite, it has been shown in Bakos and Brynjolfsson [10] that the grand bundle selling is fully revenue extracting, as N goes to infinity. Hence grand bundle selling provide a natural upper bound for asymptotic analysis. On the other hand, separate selling is also important because the loot boxes may be subject to regulation, and grand bundle with huge size and expensive price is unrealistic in practice.

Customer Behaviour

In order to have a sensible model for loot box selling, one must model how a customer values the random allocation, and how a customer valuation for the loot box adjusts after multiple purchases. Four factors affect the customer purchase behavior: the pre-announced probabilistic allocation, the valuation function of the box, the customer decision strategy, and the realized sample paths. As mentioned previously, the probability allocation is uniform, and this fact is known to the customers. We further assume that customer is risk-neutral, meaning they value a loot box at its expected value. For the two loot box types, the utility for a loot box at price p is:

$$\text{(Unique Box)} \quad U_t = \frac{\sum_{i \in [N] \setminus S_t} V_i}{N - |S_t|} - p \quad \text{(Traditional Box)} \quad U_t = \frac{\sum_{i \in [N] \setminus S_t} V_i}{N} - p$$

Given these utilities, the customers will follow some strategy to decide when to purchase. In theory, the customer may solve a backward Bellman equation to make optimal decision that maximizes their cumulative expected utility. However, this is impractical and unrealistic for general customers as the state space increases combinatorially in the number of items. Instead, we make the natural modeling assumption that customers are myopic, i.e., they purchase if and only if their expected utility for a loot box is non-negative. We note, perhaps somewhat surprisingly, that the myopic purchasing behavior is an assumption of our model and not necessarily the optimal strategy for maximizing expected utility. In particular, when considering unique loot boxes there are scenarios in which a rational customer should purchase a loot box even if their expected utility for the purchase is negative. The following example Example 3.2.1 below demonstrates such scenario.

Example 3.2.1. *Suppose that $N = 2$, and the price of the unique box is $p = 1.6$. Further suppose the customers valuation for each item is drawn from a two-point distribution, which is either 1 or 2 with probability $\frac{1}{2}$. Consider a myopic customer*

with valuation profile $(V_1, V_2) = (1, 2)$ or $(2, 1)$. Such a customer will not buy the first box, because their expected utility for a purchase $\frac{1+2}{2} - 1.6$ is negative. However there is a strategy by which the customer with valuation profile $(1, 2)$ or $(2, 1)$ can garner positive utility in expectation. Buy the first box. If he receives the item with valuation 2, stop. Otherwise buy the second box and receive the other item. With probability $\frac{1}{2}$, the customer will receive an item valued at 2, at which point he stops and garners utility $2 - 1.6 = 0.4$. If instead he gets an item that value at 1, he then purchases the second unique box, which is then guaranteed to allocate the item they value at 2. In this second case the net utility loss is only $1+2-2*1.6 = -0.2$. The total expected utility under this strategy is $0.5 * 0.4 - 0.5 * 0.2 = 0.1 > 0$, thus the customer can gain in expectation even if their expected utility their first loot box purchase is negative.

Fortunately, we will show that as the number of possible items N tends to infinity, the normalized expected utility loss suffered by a random customer following a myopic strategy instead of an optimal one tends to zero. Further, myopic behavior is always optimal for a customer considering traditional boxes.

Theorem 3.2.1 (Myopic Purchasing Behaviour is Nearly Optimal). *Suppose valuations are drawn i.i.d. for N goods from a distribution F , and loot box strategies use a fixed price p . Then for any F and p :*

- a) *For unique loot box selling, the myopic purchasing policy is asymptotically optimal i.e, the average net utility under the myopic strategy converges to the average net utility of the optimal strategy as $N \rightarrow \infty$.*
- b) *For traditional loot box selling, the myopic purchasing policy is optimal.*

Due to the complexity/impracticality of computing the customers optimal purchasing policy, and the near optimality of the myopic purchasing rule, we believe Theorem 3.2.1 provides compelling evidence that restricting to myopic purchasing

behaviour does not degrade the predictive power of our models. For the remainder of this paper we will assume customers behave myopically. However, theorems Theorem 3.3.1(a) and Corollary 3.4.1, regarding lower bounds on the revenue of unique box are also valid for strategic customers. This is because customers following a myopic strategy leads to the most conservative estimation of the selling volume i.e. the number of loot boxes purchased. In the cases where the optimal purchasing strategy differs from the myopic strategy, that customer purchases *more* loot boxes. Thus the revenue of a loot box strategy under the assumption of myopic customer behaviour is a lower bound on the revenue when customers purchase optimally.

Relations between (UB), (TB), (SS) and (GB) for finite N .

In this work we aim to understand when loot boxes are an effective revenue management tool. In particular we would like to establish relations between the revenues of (UB), (TB), (SS), and (GB). Unfortunately when the number of items N is fixed, only a single relationship exists for all valuation distributions. Specifically, a unique loot box strategy can never exceed the revenue of a grand bundle strategy.

Observation 3.2.2. $\mathcal{R}_{GB} \geq \mathcal{R}_{UB}$ for any N and distribution $V \sim F$.

This is because the condition for purchasing a grand bundle at price Np is the same as the condition for purchasing the first unique box with price p i.e. $\frac{\sum_i V_i}{N} \geq p \equiv \sum_i V_i \geq Np$. However, under a unique box strategy it may be the case that the customer ends up purchasing less than N loot boxes which implies the relation. This relation does not extend to grand bundle selling and traditional loot box selling. The condition for purchasing the first traditional box remains the same, however in this case a customer may buy multiple traditional boxes before obtaining a new item. In fact, for any other proposed order relation between two of the four strategies, there exists a distribution for which it holds and vice versa. We list the nine remaining

possible revenue relations along with distributions for which these order relation holds in Table 3.1. These distributions were found by numerically searching over the space of two-point and three-point valuation distributions, and enumerating the possible prices and customer valuations.

Table 3.1: Possible Relations Between (UB) , (TB) , (GB) and (SS) .

Relation	N	Valuation
$GB > UB > TB > SS$	3	$P(V_i = 0.98) = 1/2, P(V_i = 2.02) = 1/6, P(V_i = 3.01) = 1/3$
$GB > UB > SS > TB$	10	$P(V_i = 1) = P(V_i = 2.75) = 1/2$
$GB > TB > UB > SS$	3	$P(V_i = 1.01) = 1/2, P(V_i = 1.98) = 1/6, P(V_i = 3.03) = 1/3$
$GB > TB > SS > UB$	2	$P(V_i = 1) = 1/2, P(V_i = 2.02) = 1/6, P(V_i = 3) = 1/3$
$GB > SS > UB > TB$	2	$P(V_i = 1) = P(V_i = 2.3) = 1/2$
$GB > SS > TB > UB$	2	$P(V_i = 1) = P(V_i = 2.75) = 1/2$
$SS > TB > GB > UB$	2	$P(V_i = 1) = P(V_i = 100) = 1/2$
$SS > GB > UB > TB$	10	$P(V_i = 1) = 3/10, P(V_i = 10) = 7/10$
$SS > GB > TB > UB$	4	$P(V_i = 1) = 3/10, P(V_i = 10) = 7/10$

Motivated by the inconsistency of the four strategies when N is finite, for the rest of this paper we will focus on asymptotic analysis. As it turns out, in an asymptotic regime a consistent ordering emerges.

3.3 Asymptotic Analysis of Loot Box Pricing

In this section we will study loot box strategies in an asymptotic regime. The need for asymptotic analysis is justified by Theorem 3.2.1, and by the general incomparability of the various selling strategies in the scope of this study, see Section 3.2. Further, an asymptotic regime is well motivated in practice where N , the number of items sold in the video game, is often in the thousands. For example, in the popular online games *Dota 2* or *Overwatch*, the number of cosmetic items sold through loot boxes exceeds 3500.

In this asymptotic regime we will show that unique box strategies earn normalized revenues of μ and traditional box strategies earn normalized revenues of $\frac{\mu}{e}$. Since the

expected normalized revenue of any selling strategy cannot exceed μ , this result proves that unique box and traditional box strategies are asymptotically optimal and sub-optimal, respectively. Further, we can compare the performance of these two loot box strategies with the performance of grand bundle selling and separate selling in this regime. By an easy application of the strong law of large numbers, it is known that grand bundle also obtains normalized revenue of μ (see [10] for a detailed discussion of grand bundle strategies). On the other hand, the revenue of separate selling strategies depends explicitly on the distribution of customer valuations, and can earn anywhere between 0% and 100% of the normalized revenue.

Theorem 3.3.1 (Asymptotic Revenue of (UB), (TB)). *Suppose valuations are drawn i.i.d. for N goods from a distribution F with finite mean μ , and variance σ^2 .*

a) *Then unique loot box selling strategies are guaranteed to earn,*

$$\mathcal{R}_{UB} \geq \mu(1 - N^{-1/5}) \left(1 - (1 + \frac{2\sigma^2}{\mu^2})N^{-1/5} - \frac{\sigma^2}{\mu^2}N^{-3/5} + \frac{\sigma^2}{\mu^2}N^{-4/5} \right).$$

b) *Then traditional loot box selling strategies are guaranteed to earn,*

$$\begin{aligned} \mathcal{R}_{TB} &\geq \frac{\mu}{e} \log \left(\frac{1}{\frac{1}{e} + \frac{1 + \frac{\sigma^2}{\mu^2}}{N}} \right), \\ \mathcal{R}_{TB} &\leq \bar{p}_N \left(\zeta_N - \log(1 - N^{-\frac{1}{3}}) \right) + \bar{p}_N \log \frac{\mu}{\bar{p}_N} + \frac{(1 - N^{-\frac{1}{3}})\sigma^2 \log N}{\mu N^{\frac{1}{3}}}, \end{aligned}$$

where γ is the Euler-Mascheroni constant, $\zeta_N = \sum_{i=1}^N \frac{1}{i} - \log(N) - \gamma$, and $\bar{p}_N = \mu / \exp(1 + \log(1 - N^{-\frac{1}{3}}) - \zeta_N)$.

Moreover, letting N tend to infinity we have:

$$\lim_{N \rightarrow \infty} \mathcal{R}_{UB} = \mu \qquad \lim_{N \rightarrow \infty} \mathcal{R}_{TB} = \frac{\mu}{e}.$$

This result has a number of important implications for a monopolist considering loot box strategies. First, Theorem 3.3.1 highlights an important design aspect of loot boxes, namely that the ability to monitor a customers current inventory and appropriately control their allocation can hugely increase revenue. With information of a customers inventory, a seller can implement unique boxes which are asymptotically revenue optimal and enjoy a host of additional benefits (see Section 3.4). With this information, the seller is restricted to traditional loot box designs that, while guaranteed to garner a constant fraction of the revenue, are also fixed to that constant, always earning only $\frac{1}{e}$ of the optimal revenue regardless of the valuation distribution. Together, Theorem 3.3.1 yields a compelling rebuttal to the seemingly sound wisdom that traditional boxes induce enough additional purchases to offset their inherently lower prices. As we show, it is in fact the opposite. In an asymptotic regime with optimally chosen prices, both loot box strategies will induce the *same number* of expected purchases. Specifically, in expectation under the optimal prices, a customer will purchase traditional boxes until they have collected $1 - \frac{1}{e}$ fraction of the catalog of items, requiring on average $\frac{e}{e-1}$ purchases per item. On the other hand, when faced with optimally priced unique boxes, a customer will collect almost all the items. Thus the expected normalized number of loot box purchases under optimal prices are equal.

We also emphasize that the revenue guarantees provided by Theorem 3.3.1 are based on explicit prices for the loot boxes, namely $p = \mu(1 - \frac{1}{N^{1/5}})$ for unique boxes, and $p = \frac{\mu}{e}$ for traditional boxes, respectively. It is interesting to note that these prices do not depend on the customers actual valuation distributions, except through dependence on the mean. Such a guarantee is called a *distribution-free* ([33]) bound on the revenue, since it does not depend on the distributions themselves but merely statistics of them. The distribution free nature of the lower bounds from Theorem 3.3.1 imply that the revenue guarantee extends to the case where customers have heterogeneous

valuation distributions, and assures that simultaneously all customers types are well handled by a common loot box price.

Further, we can compare the rates of convergence in Theorem 3.3.1 against the known rates of convergence for grand bundle selling. For unique boxes, the rate at which the expected revenue tends to μ is $O(N^{-1/5})$, whereas for grand bundle selling, the convergence rate to μ is $\Theta(N^{-1/3})$ (as shown in [10]). While it is not surprising that unique box strategies tend to μ slower (recall by Observation 2 the revenue of a grand bundle always dominates unique box for every N), it is interesting that their convergences are comparable even when unique boxes must be bought one at a time. Further, if the valuations V_i are bounded, a reasonable assumption for a virtual item in a video game, then one can easily strengthen the convergence rates for unique box strategies using Chernoff’s inequalities. When valuations are bounded one can achieve sub-Gaussian convergence rate to μ for both unique box selling and grand bundle selling.

In light of Theorem 3.3.1, it is worthwhile to discuss why traditional loot boxes are popular, given their substantially lower expected revenue. We propose a number of possible explanations. First, traditional boxes such as Gachapon and Pokemon card packs exist even before the digital age (and subsequently video games), and may continue as a hold over from those times. Second, traditional boxes may be easier for customers to understand but intuitively, and from the perspective of Theorem 3.2.1 in which we showed the simple myopic policy is optimal. Third, when N is small, it is possible that traditional box selling has relatively good performance (c.f. Table 3.1). Fourth, our analysis assumes an unlimited budget. When there is a finite budget, unique box, traditional box and separate selling all have the same expected revenue since all three strategies extract all the budget in the limit. Fifth, the behaviour of possibly irrational customer not captured in our model may have a large impact on revenue. It is possible that a fraction of customers, so called *whales*, may wish to

collect all of the items regardless of cost (e.g., to flaunt their collection or to compete with other customers), and hence violates the risk-neutral assumption. In this case, the revenue from traditional boxes may increase dramatically, and outperform other selling strategies. Lastly, the presence of salvage systems may increase the revenue of a traditional box strategy, we will study this possibility in detail in Section 3.4.

Finally, we discuss the insights on customers' surplus. From a customers perspective, one might naturally assume that since unique boxes always allocate unique items, it may benefit them when compared to traditional boxes. We show this assumption is unfounded. It follows from Theorem 3.3.1 that a profit maximizing monopolist using unique boxes can garner all utility. On the other hand, under traditional boxes the limiting normalized net utility it is $((1 - 1/e)N\mu - N\mu/e) / N = (1 - 2/e)\mu$, thus even when prices are optimally chosen a customer still obtains positive expected utility. Hence a customer is actually worse off when facing unique boxes over traditional boxes.

3.4 Loot Box Design Problem

In the previous section we studied the loot box pricing problem in a simplified model. Recall our loot box models assumed that the valuations for all items were i.i.d., that the probability an item is allocated by a loot box is uniformly random, that all customers are myopic utility maximizers, and that customers obtain no value from duplicate items. In practice many of these assumptions are violated, in this section we extend the results of the previous section to these cases, and derive insight into how loot box strategies should be designed, and increasing the applicability of our results.

Joint Allocation and Pricing for Multiple Classes of Items

In the previous section, we assumed that valuations for all items were i.i.d., and that each item (*unowned* item in the case of unique boxes) was equally likely to be allocated by the loot box. In practice, both these assumptions may be grossly inaccurate. Often in online games the items are explicitly grouped based on rarity or effectiveness, e.g., in the online popular game *PlayerUnknown's Battlegrounds*, customers may get Mythic, Legendary, Epic, or Rare items from a loot box (see Fig. 3.4). In these cases items of the same group may be represented by draws from the same distribution, but between groups the items value will differ wildly. In this subsection we will extend our model to allow for M *classes* of items, where each item in class $m \in [M]$ is drawn from a distribution $V^m \sim F_m$.

The introduction of multiple item classes allows for some items to significantly more valuable than others, it is then possible in this case that non-uniform allocation probabilities may be reasonable. We will denote the proportion of the items in each class m as β_m , where $\sum_{m \in [M]} \beta_m = 1$. For asymptotic results we will suppose each the number of items in each class grows proportionally to N i.e. $\beta_m N$. Further, we will allow the allocation probabilities for each class of items by the loot box to be determined by the seller. A loot box strategy is now characterized by a price p and a set of allocation probabilities. Our goal is to characterize the revenue optimal combination of price and allocation probabilities for loot boxes over multiple classes of items.

For unique box strategies, the optimal non-uniform allocation probabilities are dynamic and depend on the customers current set of items. It is thus difficult to explain such policies to customers, let alone characterize the optimal allocation probabilities. Thankfully, for unique boxes the most natural extension to the multi-class case is asymptotically optimal. We will define a *uniform* unique box strategy as one which allocates items by choosing an item which the customer does not yet own uniformly

at random, regardless of class. Note that this is the same allocation rule as in Theorem 3.3.1, each item initially appears in the loot box with probability $\frac{1}{N}$. For uniform unique boxes, the analysis of the allocation probabilities and prices is straightforward. When $M \ll N$ one can simply apply the analysis of Theorem 3.3.1(a) with a slight modification. We encapsulate this observation in the following corollary.

Corollary 3.4.1 ((UB) with Uniform Allocations are Asymptotically Optimal). *Suppose there are M classes of items, and valuations in class $m \in [M]$ are drawn i.i.d for $\beta_m N$ goods from a distribution F_m with mean μ_m and variance σ_m . Let $\bar{\mu} = \sum_m \beta_m \mu_m$ and $\bar{\sigma}^2 = \sum_m \beta_m \sigma_m^2$. Then,*

$$\mathcal{R}_{UB} \geq \bar{\mu}(1 - N^{-1/5}) \left(1 - \left(1 + \frac{2\bar{\sigma}^2}{\bar{\mu}^2} \right) N^{-1/5} - \frac{\bar{\sigma}^2}{\bar{\mu}^2} N^{-3/5} + \frac{\bar{\sigma}^2}{\bar{\mu}^2} N^{-4/5} \right),$$

Moreover,

$$\lim_{N \rightarrow \infty} \mathcal{R}_{UB} = \bar{\mu}.$$



Crate Odds		
?	Mythic Item	0.3%
?	Legendary Item	1.0%
?	Epic Item	22.7%
?	Rare Item	76%

Figure 3.4: In the game *PlayerUnknown's Battlegrounds*, the traditional box contains four classes of items: Mythic, Legendary, Epic, and Rare. The allocation probability for items of different classes varies, however items within the same class have the same probability.

Our main focus in this section is to understand how non-uniform allocation rules affect traditional box strategies. The most pressing question in this vein is whether

or not the additional flexibility in allocation rules can increase the revenue of traditional box strategies beyond the $\bar{\mu}/e$ guarantee of Theorem 3.3.1(b). Let $\{q_i\}_{i=1}^N$ be a strategies allocation probabilities, where q_i is the probability of allocating item i , $\sum_{i=1}^N q_i = 1$. Fixing a price p , the *allocation problem* is to find the best vector of probabilities $\mathbf{q}^* \in \mathcal{Q} = \{q_i | \sum_{i=1}^N q_i = 1\}$ that maximizes the revenue. The joint pricing and allocation problem is to find the optimal price p^* and allocation probabilities \mathbf{q}^* which yield the maximum revenue. To simplify the problem, we will restrict our attention to class level allocation probabilities, i.e., allocation rules where all items in the same class have the same allocation probabilities. We emphasize that class level allocation rules are common in practice (e.g., Figure 3.4).

For a class level allocation rule, we will let $d_m = \sum_{j \in \text{Class } m} q_j = \beta_m N q_j$ be the probability of allocating an item in class m , and $\mathbf{d} = (d_1, \dots, d_M)$. Further, let $Q_{\mathbf{d}}^N(p)$ and $\mathcal{R}_{TB}(p, \mathbf{d})$ be the normalized number of loot boxes purchased by a customer (i.e. $\frac{\mathbb{E}[\# \text{ Purchases}]}{N}$), and the normalized expected revenue, under price p and class allocation probabilities \mathbf{d} , respectively. To enable our study of traditional boxes with non-uniform allocation rules we first show useful relationship between price, allocation probability and the limiting selling volume.

Lemma 3.4.1. *Suppose there are M classes of items, and valuations in class $m \in [M]$ are drawn i.i.d for $\beta_m N$ goods from a distribution F_m with mean μ_m and variance σ_m . Suppose a traditional box strategy follows a multi-class allocation rule \mathbf{d} , and price $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$, for some $k \geq 0$. Then*

$$\lim_{N \rightarrow \infty} \mathbb{E} [Q_{\mathbf{d}}^N(p)] = k.$$

Armed with Lemma 3.4.1, we find that as N tends to infinity, the optimal solution for the joint pricing and multi-class allocation problem has a surprisingly simple structure. Namely, proportional multi-class allocation probabilities with the same

prices as in Theorem 3.3.1(b) garners more revenue than any other pricing and allocation rule, and again the normalized revenue converges to $\frac{\bar{\mu}}{e}$. Thus answering the question of whether exotic allocation probabilities can greatly increase the revenue of traditional boxes firmly in the negative.

Theorem 3.4.1 ((TB) with Proportional Allocations are Asymptotically Optimal).

Suppose there are M classes of items, and valuations in class $i \in [M]$ are drawn i.i.d for $\beta_i N$ goods from a distribution F_i with mean μ_i and variance σ_i . Let $\bar{\mu} = \sum_i \beta_i \mu_i$ and $\bar{\sigma}^2 = \sum_i \beta_i \sigma_i^2$. Then in the limit, we have

$$\max_{p, \mathbf{d}} \lim_{N \rightarrow \infty} \mathcal{R}_{TB}(p, \mathbf{d}) = \frac{\bar{\mu}}{e} \quad \text{and} \quad \arg \max_{N \rightarrow \infty} \lim_{N \rightarrow \infty} \mathcal{R}_{TB}(p, \mathbf{d}) = \left(\frac{\bar{\mu}}{e}, (\beta_1, \dots, \beta_M) \right).$$

Theorem 3.4.1 provides a natural generalization of Theorem 3.3.1(b) to the multi-class case. The proportional allocation strategy with price $\bar{\mu}/e$ is asymptotically optimal, for any (finite) number of classes, and with any class-wise distributions. In this sense Theorem 3.4.1 makes a traditional loot box sellers decision simple, instead of designing complicated allocation structures, simply use proportional allocations and focus on the price. Further, Theorem 3.4.1 extends the asymptotic dominance of unique box strategies over traditional boxes to the case of multiple item classes; varying the allocation probabilities cannot close the gap in revenue between the two strategies.

Finally, while Theorem 3.4.1 provides a simple solution to the joint allocation and pricing problem for traditional loot boxes, this simplicity depends critically on the seller using the revenue optimal price. However there are situations where a seller may use a price that differs from the theoretical optimum. Market pressures, platform stipulations, promotional rounding (i.e. the “optimal” price may be \$1.07, which is then rounded to \$0.99), and other factors may entice sellers to offer a price other than $\bar{\mu}/e$. In these cases, when the price is fixed and sub-optimal, the optimal

class allocation probabilities may not be proportional. In such cases we may lean on Lemma 3.4.1 to compute nearly optimal allocation probabilities. To do so, suppose the target selling volume kN is fixed and exogenous, then the maximum possible price (and thus revenue, since volume is fixed) which achieves this selling volume can be solved for by Lemma 3.4.1 via the following mathematical program,

$$\begin{aligned}
(OPT_{kN}) = \max \quad & \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k} \\
\text{s.t.} \quad & \sum_{m=1}^M d_m = 1 \\
& d_m \geq 0, \forall m.
\end{aligned} \tag{3.1}$$

The objective function for (OPT_{kN}) is quasi-concave, and the constraints are linear, thus it can be solved efficiently (see [77]). A seller may then search for the maximum selling volume k by performing exponential search on $[0, \infty)$, and solving (3.1) at each iteration to see if k is feasible to the given price.

Multi-item Loot Boxes

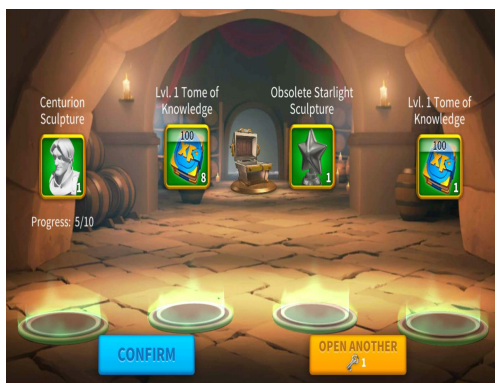
In the previous section we assumed each loot box allocated only a single item, however in many games and almost all offline versions of traditional boxes (for example packs of cards), the loot box contains multiple items. Figs. 3.5a and 3.5b shows examples of size- j boxes in practice. In this section we will show that Theorem 3.3.1 easily extends to the case where loot boxes are of sizes larger than 1 by imagining the sequence of valuations as a sequence of k fold convolutions of the random variables. That is, by replacing the sequence of valuations for items allocated, $\{V_i\}_{i=1}^N$, by $\{\sum_{j=ki}^{ki+k} V_j\}_{i=1}^\infty$ and applying Theorem 3.3.1. Further, by independence, the mean and variance of the k -fold convolution is $\mu_k = k\mu$, and $\sigma_k = k\sigma$ respectively and the the same rates of convergence apply with $N_k = N/k$. We formalize this observation in the following

corollary.

Corollary 3.4.2 (Multi-Item Loot Boxes). *Suppose valuations are drawn i.i.d. for N goods from a distribution F with finite mean μ , and variance σ^2 , and each loot box is of some fixed size k . Then,*

$$\lim_{N \rightarrow \infty} \mathcal{R}_{UB} = \mu$$

$$\lim_{N \rightarrow \infty} \mathcal{R}_{TB} = \frac{\mu}{e}.$$



(a) Size 4 Box

WHAT'S IN A HEXTECH CHEST?	
Content Type	% of Openings
Skin Shard	50%
Champion Shard	25%
Emote Permanent	10%
Ward Skin Shard + 150 Bonus Orange Essence	11.5%
Summoner Icon Shard + 150 Bonus Orange Essence	3.5%
BONUS DROPS	
Gemstone	3.6%
Bonus Chest + Key	10%

(b) Random Size Box

Figure 3.5: The left panel shows an implementation of a loot box in the mobile game *Rise of the King*. Each box outputs 4 items. The right panel shows an implementation of a loot box in online game *League of Legends*. Customers are guaranteed to get one item, with the chance to get at most 2 bonus items within one box.

Salvage Costs

In previous sections we assumed that customers had no value for duplicate items, and that a customer could not trade in low valued items back to the seller. In practice however, many loot box marketplaces are equipped with *salvage* systems, mechanisms by which a customer can trade in unwanted items for currency. Salvage systems are a ubiquitous method for managing customer satisfaction under loot box policies, offering customers a form of recourse against unlikely or unfortunate outcomes, and mitigating worst case allocations from the loot box. In Fig. 3.6 we provide one example. In this section we will consider loot box selling strategies that allow customers

to trade in or *salvage* items for some amount of virtual currency c . For simplicity we will restrict our attention to the case when the number of items in the loot box is 1 and there is only a single class of items; extensions beyond these assumptions follow from applying the analysis in Sections 3.4 and 3.4. We will also assume that when a customer receives an item they value at less than c , they will *immediately salvage the item* before continuing to purchased new loot boxes.



Figure 3.6: In the game *Dota 2*, players can trade in 6 unwanted items for a new loot box, plus 2000 shards, a form of in-game currency.

The main insight will be in understanding the two competing affects that introducing salvage systems have on loot box revenue. On the one hand, the presence of a salvage cost c increases the minimum valuation of any item to at least c , increasing the expected valuations (from $\mathbb{E}[V]$ to $\mathbb{E}[\max\{V, c\}]$) and inducing more purchases. On the other hand, salvage systems return currency to the customer diluting the revenue garnered from customer purchases. The main results in this section will be characterize and extend the revenue guarantees of Theorem 3.3.1 to the case when items can be salvaged for some cost c .

We will use the notation \mathcal{R}^c to denote the optimal revenue of a strategy with fixed salvage cost c . Note that in the presence of a salvage system, the allocation mechanism for unique box strategies is no longer well specified. Specifically, should unique box strategies allocate items that a customer does not currently own, or items that a customer has never owned? That is, should items that a customers has salvaged

previously be allowed to be allocated to a customer again. For our results we will assume the latter, that customers are never allocated an item they had previously salvaged.

The additional richness in strategies introduced by considering salvage systems also allows us to pin down the precise relationships between loot box strategies and separate selling. Specifically, by treating the salvage cost c as a parameter of a loot box strategy, the revenue of either an optimal unique box or traditional box strategy can be guaranteed to dominate the revenue of separate selling.

Proposition 3.4.1 (Loot Boxes with Salvage Outperform Separate Selling). *For any N ,*

$$\max_c \mathcal{R}_{UB}^c \geq \mathcal{R}_{SS} \text{ and } \max_c \mathcal{R}_{TB}^c \geq \mathcal{R}_{SS}.$$

Note that this result is valid for finite N . It is well known that grand bundle is not guaranteed to outperform separate selling for finite N , even though grand bundle selling is asymptotically optimal. This further explains the power and popularity of loot box in practice.

We now investigate the revenue of salvage systems in the asymptotic scheme. The following theorem gives the limiting revenue with respect to a fixed c .

Theorem 3.4.2 (Revenue of Loot Box Strategies with Salvage Costs.). *Suppose valuations are drawn i.i.d. for N goods from a distribution $V \sim F$, with mean μ and variance σ^2 . Let c be the salvage cost, $\gamma = F(c)$ be the probability of salvage, and $\eta = \mathbb{E}[V|V > c]$.*

a) *Unique loot box selling strategies asymptotically earn normalized revenue,*

$$\lim_{N \rightarrow \infty} \mathcal{R}_{UB}^c = \eta(1 - \gamma).$$

b) *Traditional loot box selling strategies asymptotically earn normalized revenue,*

$$\lim_{N \rightarrow \infty} \mathcal{R}_{TB}^c = (1 - \gamma)(\eta - c) \left(\frac{c}{\eta - c} + e^{-\frac{\eta}{\eta - c}} \right).$$

This result carries through a number of the same insights from Theorem 3.3.1 to the case with salvage costs. First, note that like before the (asymptotic) revenue of unique box strategies dominates the (asymptotic) revenue of traditional box strategies for any valuation distribution F and salvage cost c . To see this note since $0 < c \leq \eta$, we may substitute c by $c = q\eta$, for some $q \in [0, 1]$. Plugging in this substitution and rearranging yields the upper bound,

$$\lim_{N \rightarrow \infty} \mathcal{R}_{TB}^c \leq (1 - \gamma)\eta \max_{q \in (0, 1]} \left(q + (1 - q)e^{-\frac{1}{1-q}} \right) \leq (1 - \gamma)\eta, \quad (3.2)$$

where the final equality comes from noting $q + (1 - q)e^{-\frac{1}{1-q}}$ is monotone increasing and tends to 1 as $q \rightarrow 1$. Thus $\lim_{N \rightarrow \infty} \mathcal{R}_{UB}^c \geq \lim_{N \rightarrow \infty} \mathcal{R}_{TB}^c$. Further, the monotonicity in the maximum in Eq. (3.2) implies when $\frac{c}{\eta}$ is large (close to 1), the gap in expected revenue between unique box strategies and traditional box strategies is small and generally shrinks from a factor of e ($\frac{c}{\eta} = 0$) monotonically down to 1 ($\frac{c}{\eta} = 1$). Thus when salvage costs are relatively large, the additional value of employing unique box strategies decreases. Moreover, note that for unique boxes, $(1 - \gamma)\eta \leq \mu$, with equality achieved only when $c = 0$. Lastly, by combining Proposition 3.4.1 and Theorem 3.4.2, we can have a complete order of the four strategies in the limit, in presence of salvage cost:

$$\lim_{N \rightarrow \infty} \mathcal{R}_{GB} = \lim_{N \rightarrow \infty} \max_c \mathcal{R}_{UB}^c \geq \lim_{N \rightarrow \infty} \max_c \mathcal{R}_{TB}^c \geq \mathcal{R}_{SS}.$$

Finally, recall salvage systems are primarily used to improve customer outcomes and overall satisfaction with the system. Formally, the expected normalized customer surplus under a revenue maximizing unique box with salvage cost c is $(1 - \gamma)\eta + \gamma c -$

$(1 - \gamma)\eta = c\gamma$, since customers will purchase almost all goods. Thus the expected normalized customer surplus is monotonically increasing in c . The salvage system thus enables the seller to balance allocate the revenue and customer surplus to their desired proportion.

3.5 Numerical Tests

In this section we conduct numerical experiments to demonstrate the efficacy of unique box and traditional box selling in the large but finite N regime. In particular, for a set of typical valuation distributions, we will compute and compare the revenue of (UB), (TB), (GB), and (SS) strategies as N increases. For (UB) and (TB) strategies we will also use the heuristic prices motivated by Theorem 3.3.1, specifically for (UB) we will use price $\mu(1 - N^{-1/5})$ and for (TB) we will use the price $\frac{\mu}{e}$. Note the (TB) price differs from the price described in Theorem 3.3.1(b), we instead use $\frac{\mu}{e}$ due to the overly poor performance of the rule in the theorem when N is small, and because the price in the theorem quickly converges to $\frac{\mu}{e}$ anyway. Further, numerically we demonstrate that the instance optimal prices quickly converge to these heuristic prices.

In our experiments we will assume the customers valuation is drawn from one of three distributions: Uniform between 0 and 2, Log-normal with log-mean 0 and log-variance 1, and Exponential with mean 1. These distributions are commonly used to model customer valuation, and have been previously studied in [1]. Under each valuation distribution, and for N ranging from 3 to 3000, we will consider the revenue of 7 selling strategies: (i) traditional box with optimal price (TB Optimal), (ii) traditional box with price μ/e (TB Heuristic), (iii) unique box with optimal price (UB Optimal), (iv) unique box with price $\mu(1 - N^{-1/5})$ (UB Heuristic 1), (v) unique box with the optimal price of grand bundle (UB Heuristic 2), (vi) grand bundle selling

(GB) and (vii) separate selling (SS). Computation of the revenues will be done via simulation, by generating 50000 customer sample paths and using brute force to search the optimal prices (at 0.01 accuracy).

In Figure 3.7, we plot how the optimal price changes as N increases. Note the optimal price of a traditional box strategy quickly converges to μ/e , and is nearly indistinguishable for $N \geq 50$ while the optimal price of unique box strategy converges to μ , albeit somewhat slower. Further, the optimal price for unique box strategies are much higher than $\mu(1 - N^{-1/5})$, indicting a more aggressive pricing heuristic may perform better than the overly conservative heuristic prices we used to construct a lower bound. Interestingly, the optimal prices for unique box strategies and grand bundle strategies are quite similar, lending further evidence that (UB Heuristic 2) may be an effective strategy in practice.

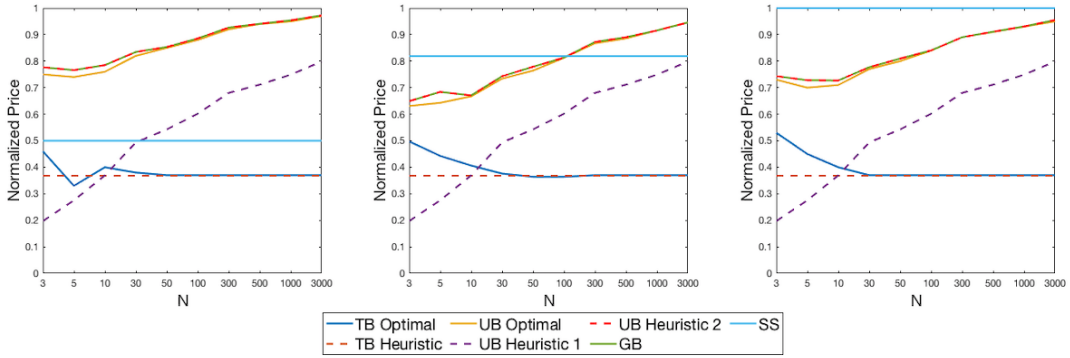


Figure 3.7: Normalized Prices for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.

In Figure 3.8, we plot the normalized revenues for each strategy. For traditional boxes, both the optimal and heuristic policy quickly converge μ/e . For unique boxes, although both heuristic strategies earn less than the optimal grand bundle strategy, they appear quite effective, in all three cases garnering 70% of the maximum revenue when $N \geq 100$. We can also see that unique box with grand bundle price (UB Heuristic 2) gives almost the same revenue as the optimal price, while the less aggressively priced strategy (UB Heuristic 1) lags markedly behind. Recall the intuition

behind (UB Heuristic 1) in Theorem 3.3.1 was to use a price and always maintained the lower bound of normalized selling volume that was close to 1. As it turns out, (UB Heuristic 1) is too conservative, because in most cases, the selling volume will be much higher than the theoretical lower bound. Finally, we emphasize that in practice (UB Heuristic 2) is easy to implement, because computing the optimal price of grand bundle is computational less difficult than that of unique box.

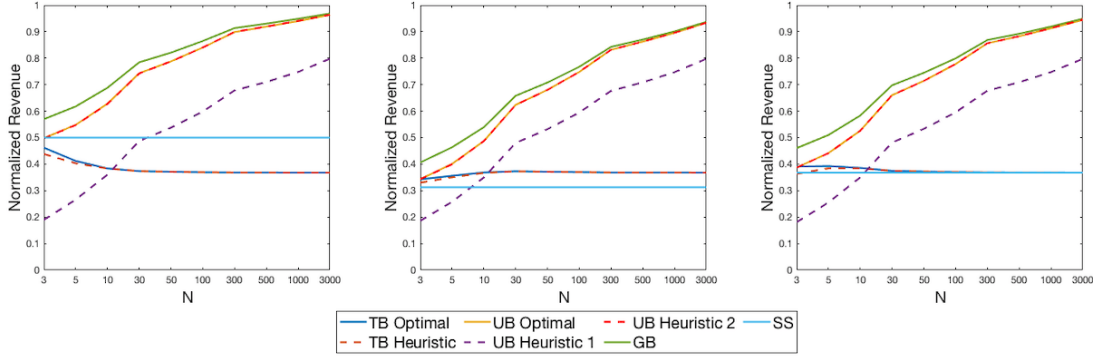


Figure 3.8: Normalized Revenue for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.

In Figure 3.9, we show the normalized surplus for different strategies. While the surplus of traditional box and separate selling converges to a constant, the surplus of unique box and grand bundle converges to 0.

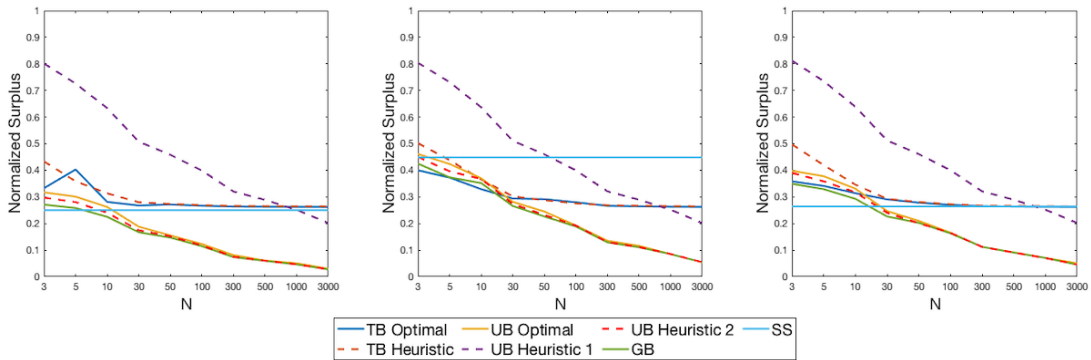


Figure 3.9: Normalized Surplus for Uniform (left), Log-normal (middle) and Exponential (right) Valuations.

3.6 Conclusions

In this paper we analyze the revenue of loot box selling strategies, which are among the popular, profitable, and controversial selling strategies in the video game industry. We first show that myopic behavior is reasonable for customers, and investigate pricing and design problems under this assumption. Motivated by the generally incompatibility in the finite N setting, we shifted our attention to an asymptotic regime where we then proved that the unique loot box strategies are fully revenue extracting, whereas the traditional box design can garner only a fixed factor, $1/e$, of the maximum revenue. We then extended our framework to analyze three aspects of loot box design. First, we considered the case of multiple item classes, and showed that simple proportional allocations with an appropriate price are asymptotically optimal under any risk-preference. Second we show that our results can be extended to multi-item boxes. Finally, we considered salvage systems and showed they may further increase revenue, while simultaneously enjoying customer friendly guarantees.

Our work implies a host of managerial insights for sellers, customers, and regulators of loot boxes. For sellers, we give a thorough analysis for the profitability of loot boxes and encouraging numerical results, yielding guidelines for how to design and price loot boxes so as to maximize revenue. For customers, we confirm that myopic purchasing behavior is a nearly optimal purchasing policy. Further, we clarify many common misconceptions about loot boxes. Namely, that unique boxes cost more and when optimally priced, leave less surplus for customers. Moreover, while customers may regard salvage systems as treaties from the seller, we show that such designs are a method to augment the revenue. Finally, for regulators, we show that although unique boxes are less random and thus inherently less like gambling, such boxes actually lower customer surplus. We also highlight the importance of supervision with regard to a sellers allocation rule, as sellers may gain extra revenue by using even a random perturbation strategy.

While the buzz around loot box selling is booming in the media, there is a distinct lack of academic work which evaluates this simple selling strategy in the broader context revenue management systems for selling multiple items sequentially. While our work breaks ground on this topic, there are number of avenues yet unexplored by our work. In particular we focuses on comparisons between loot box selling and other simple, popular mechanisms. One particularly fruitful avenue of potential research would be to study loot boxes via the lens of optimal mechanism design, with the aim of characterizing precisely under what conditions loot box selling is the best possible mechanisms, or if not, identifying better alternatives. Further, a number of natural extensions of loot box could be considered. For instance, personalized and dynamic allocation rules utilizing information about the customers current inventory. Finally, connecting with current media, it would be interesting to consider loot box pricing and design problems under more legal constraints or fairness considerations.

Bibliography

- [1] T. Abdallah. On the benefit (or cost) of large-scale bundling. *Production and Operations Management*, 0(0), 2018. doi: 10.1111/poms.12958. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.12958>.
- [2] T. Abdallah, A. Asadpour, and J. Reed. Large-scale bundle size pricing: A theoretical analysis. *Available at SSRN: <https://ssrn.com/abstract=2907118>*, 2017.
- [3] T. Abdallah, A. Asadpour, and J. Reed. Large-scale bundle size pricing: A theoretical analysis. 2017.
- [4] S. P. Anderson, A. De Palma, and J. F. Thisse. *Discrete choice theory of product differentiation*. MIT press, 1992.
- [5] Apple. App store review guidelines. 2018. URL <https://developer.apple.com/app-store/review/guidelines/>.
- [6] N. Arora, X. Dreze, A. Ghose, J. D. Hess, R. Iyengar, B. Jing, Y. Joshi, V. Kumar, N. Lurie, S. Neslin, S. Sajeesh, S. N. Su, Meng, J. Thomas, and J. Zhang. Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters*, 19(3-4):305, 2008.
- [7] G. Aydin and S. Ziya. Personalized dynamic pricing of limited inventories. *Operations Research*, 57(6):1523–1531, 2009.
- [8] P. Azar, C. Daskalakis, S. Micali, and S. M. Weinberg. Optimal and efficient parametric auctions. In *Proceedings of the twenty-fourth annual ACM-SIAM symposium on Discrete algorithms*, pages 596–604. SIAM, 2013.
- [9] M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *the Proceedings of the 2014 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, 2014.
- [10] Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits, and efficiency. *Management Science*, 45(12):1613–1630, 1999. doi: 10.1287/mnsc.45.12.1613. URL <https://doi.org/10.1287/mnsc.45.12.1613>.
- [11] F. Balestrieri and S. Izmalkov. Informed seller in a hotelling market. *SSRN*, 2016. URL <http://dx.doi.org/10.2139/ssrn.2398258>.

- [12] F. Balestrieri, S. Izmalkov, and J. Leao. The market for surprises: selling substitute goods through lotteries. 2015.
- [13] G.-Y. Ban and N. B. Keskin. Personalized dynamic pricing with machine learning. *Available at SSRN: <https://ssrn.com/abstract=2972985>*, 2017.
- [14] R. E. Barlow, A. W. Marshall, and F. Proschan. Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics*, pages 375–389, 1963.
- [15] D. Bergemann and K. Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543, 2011.
- [16] D. Bergemann and J. Valimaki. Dynamic mechanism design: An introduction. 2018.
- [17] D. Bergemann, B. Brooks, and S. Morris. The limits of price discrimination. *The American Economic Review*, 105(3):921–957, 2015.
- [18] F. Bernstein, A. G. Kök, and L. Xie. Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, 17(4): 538–553, 2015.
- [19] O. Besbes and I. Lobel. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science*, 61(1):92–110, 2015.
- [20] O. Besbes and A. Zeevi. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, 61(4):723–739, 2015.
- [21] O. Besbes, R. Phillips, and A. Zeevi. Testing the validity of a demand model: An operations perspective. *Manufacturing & Service Operations Management*, 12(1):162–183, 2010.
- [22] P. Briest and H. Roglin. The power of uncertainty: Bundle-pricing for unit-demand customers. In *Proceedings of the Workshop on Approximation and Online Algorithms (WAOA)*, pages 47–58, 2010.
- [23] P. Briest, S. Chawla, R. Kleinberg, and S. M. Weinberg. Pricing randomized allocations. In *Proceedings of the Twenty-first Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA ’10, pages 585–597, Philadelphia, PA, USA, 2010. Society for Industrial and Applied Mathematics. ISBN 978-0-898716-98-6. URL <http://dl.acm.org/citation.cfm?id=1873601.1873650>.
- [24] P. Briest, S. Chawla, R. Kleinberg, and S. Weinberg. Pricing lotteries. In *Journal of Economic Theory*, volume 156, pages 144–174, 2015.

- [25] BusinessInsider. Regulators from more than a dozen countries are looking to crack down on 'loot boxes,' a controversial video gaming practice that could be too much like gambling. 2017. URL <https://www.businessinsider.com/loot-boxes-european-regulation-2018-9>.
- [26] Y. Cai and C. Daskalakis. Extreme value theorems for optimal multidimensional pricings. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, page 522531. IEEE, 2011.
- [27] F. Caro and J. Gallien. Clearance pricing optimization for a fast-fashion retailer. *Operations Research*, 60(6):1404–1422, 2012.
- [28] L. E. Celis, G. Lewis, M. Mobius, and H. Nazerzadeh. Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science*, 60(12):2927–2948, 2014.
- [29] I. Chatzigeorgiou. Bounds on the lambert function and their application to the outage analysis of user cooperation. In *IEEE Communications Letters*, pages 1505–1508. IEEE, 2013.
- [30] S. Chawla, J. Hartline, and R. Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce (EC)*, pages 243–251, 2007. ISBN 978-1-59593-653-0.
- [31] S. Chawla, J. Hartline, D. Malec, and B. Sivan. Mutli-parameter mechanism design and sequential posted pricing. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing (STOC)*, pages 311–320, 2010. ISBN 978-1-4503-0050-6.
- [32] S. Chawla, D. Malec, and B. Sivan. The power of randomness in bayesian optimal mechanism design. In *Games and Economic Behavior*, volume 91, pages 297–317, 2015. ISBN 978-1-60558-822-3.
- [33] H. Chen, M. Hu, and G. Perakis. Distribution-free pricing. *Available at SSRN: <https://ssrn.com/abstract=3090002>*, 2017.
- [34] R. Chen, E. Gal-Or, and P. Roma. Opaque distribution channels for competing service providers: Posted price vs. name-your-own-price mechanisms. *Operations research*, 62(4):733–750, 2014.
- [35] X. Chen, I. Diakonikolas, A. Orfanou, D. Paparas, X. Sun, and M. Yannakakis. On the complexity of optimal lottery pricing and randomized mechanisms. In *the Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, 2015.
- [36] X. Chen, Z. Owen, C. Pixton, and D. Simchi-Levi. A statistical learning approach to personalization in revenue management. *Available at SSRN: <https://ssrn.com/abstract=2579462>*, 2015.

- [37] X. Chen, I. Diakonikolas, D. Paparas, X. Sun, and M. Yannakakis. The complexity of optimal multidimensional pricing for a unit-demand buyer. *Games and Economic Behavior*, 110:139 – 164, 2018. ISSN 0899-8256.
- [38] Y. Chen, S. Moorthy, and Z. J. Zhang. Research note-price discrimination after the purchase: Rebates as state-dependent discounts. *Management Science*, 51(7):1131–1140, 2005.
- [39] V. Choudhary, A. Ghose, T. Mukhopadhyay, and U. Rajan. Personalized pricing and quality differentiation. *Management Science*, 51(7):1120–1130, 2005.
- [40] S. Clifford. Shopper alert: Price may drop for you alone. *The New York Times*, Aug 2012. URL <http://www.nytimes.com/2012/08/10/business/supermarkets-try-customizing-prices-for-shoppers.html>.
- [41] M. C. Cohen, G. Perakis, and R. S. Pindyck. Pricing with limited knowledge of demand. Available at SSRN: <https://ssrn.com/abstract=2673810>, 2015.
- [42] M. C. Cohen, I. Lobel, and R. Paes Leme. Feature-based dynamic pricing. Available at SSRN: <https://ssrn.com/abstract=2737045>, 2016.
- [43] M. C. Cohen, N.-H. Z. Leung, K. Panchamgam, G. Perakis, and A. Smith. The impact of linear optimization on promotion planning. *Operations Research*, 65(2):446–468, 2017.
- [44] R. Corless, G. Gonnet, D. Hare, D. Jeffery, and D. Knuth. On the lambert w function. *Advances in Computational Mathematics*, 1993.
- [45] S. Cowan. Welfare-increasing third-degree price discrimination. *The RAND Journal of Economics*, 47(2):326–340, 2016.
- [46] D. Daskalakis and S. M. Weinberg. Symmetries and optimal multi-dimensional mechanism design. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC '12)*, 2012.
- [47] A. D’Innocenzio. Neiman marcus focuses on exclusives, personalized offers; ends merger talks. *USA Today*, Jun 2017. URL <https://www.usatoday.com/story/money/business/2017/06/13/neiman-marcus-focuses-exclusives-personalized-offers/102814962/>.
- [48] A. Drummond and J. D. Sauer. Video game loot boxes are psychologically akin to gambling. *Nature Human Behaviour*, 2(8):530–532, jun 2018. doi: 10.1038/s41562-018-0360-1. URL <https://doi.org/10.1038/s41562-018-0360-1>.
- [49] P. Dutting and M. Klimm. Revenue gaps for discriminatory and anonymous sequential posted pricing. 2016. URL [arXiv:cs.GT/1607.07105](https://arxiv.org/abs/cs.GT/1607.07105).
- [50] A. Elmachoub and Y. Wei. Retailing with opaque products. *SSRN*, 2015. URL <http://dx.doi.org/10.2139/ssrn.2659211>.

- [51] A. N. Elmachtoub and M. L. Hamilton. The power of opaque products in pricing. *Available at SSRN: <https://ssrn.com/abstract=3025944>*, 2017.
- [52] A. N. Elmachtoub and Y. Wei. Retailing with opaque products. *Available at SSRN 2659211*, 2015.
- [53] S. Fay and J. Xie. Probabilistic goods: A creative way of selling products and services. In *Marketing Science*, volume 27, pages 674–690, 2008.
- [54] S. Fay and J. Xie. Timing of product allocation: Using probabilistic selling to enhance inventory management. *Management Science*, 61(2):474–484, 2014.
- [55] K. Ferreira and J. Goh. Assortment rotation and the value of concealment. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (17-041), 2018.
- [56] J. Fingas. South Korea fines game studios over deceptive loot box odds., 2018. <https://www.engadget.com/2018/04/10/south-korea-fines-game-studios-over-loot-boxes>.
- [57] Forbes. Ea now seems legitimately terrified of loot boxes after 'battlefront 2'. 2018. URL <https://www.forbes.com/sites/insertcoin/2018/04/17/ea-now-seems-terrified-of-loot-boxes/#2d7215b7648e>.
- [58] Forbes. Report: For the first time, mobile will make up half of all gaming revenue. 2018. URL <https://www.forbes.com/sites/mattperez/2018/04/30/report-mobile-will-make-up-half-of-all-gaming-revenue/#50f486a1586a>.
- [59] G. Gallego and R. Phillips. Revenue management of flexible products. In *Manufacturing & Service Operations Management*, volume 6, pages 321–337, 2004.
- [60] G. Gallego, G. Iyengar, R. Phillips, and A. Dubey. Managing flexible products on a network. 2004.
- [61] GamesIndustry. Npd: Loot box controversy having no impact on game sales. 2017. URL <https://www.gamesindustry.biz/articles/2017-10-25-npd-loot-box-controversy-having-no-impact-on-sales>.
- [62] X. Geng. Opaque selling in congested systems. *Operations Research Letters*, 44(6):737–741, 2016.
- [63] N. Granados, K. Han, and D. Zhang. Demand and revenue impacts of an opaque channel: Evidence from the airline industry. In *Productions and Operations Management*, 2017.
- [64] A. Gronlund, K. Larsen, A. Mathiasen, J. Nielsen, S. Schneider, and M. Song. Fast exact k-means, k-medians and bregman divergence clustering in 1d. *ArXiv e-prints: arXiv:1701.07204*, 2017.

- [65] A. Gut. *Stopped random walks*. Springer, 2009.
- [66] W. Hanson and R. K. Martin. Optimal bundle pricing. *Management Science*, 36(2):155–174, 1990.
- [67] S. Hart and N. Nisan. How good are simple mechanisms for selling multiple goods? In *The Hebrew University of Jerusalem, Center for Rationality*, pages 893–906, 2014.
- [68] S. Hart and P. J. Reny. Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoretical Economics*, 10(3):893–922, 2015. doi: 10.3982/TE1517. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/TE1517>.
- [69] J. Hartline, V. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. In *Proceedings of the 17th International Conference on World Wide Web*, pages 189–198. ACM, 2008.
- [70] J. D. Hartline and V. Koltun. Near-optimal pricing in near-linear time. In F. Dehne, A. López-Ortiz, and J.-R. Sack, editors, *Algorithms and Data Structures*, pages 422–431, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg. ISBN 978-3-540-31711-1.
- [71] T. Huang and Y. Yu. Sell probabilistic goods? a behavioral explanation for opaque selling. *Marketing Science*, 33(5):743–759, 2014.
- [72] X. Huang, G. Sošić, and G. Kersten. Selling through priceline? on the impact of name-your-own-price in competitive market. *IIE Transactions*, 49(3):304–319, 2017.
- [73] A. Javanmard and H. Nazerzadeh. Dynamic pricing in high-dimensions. *Available at SSRN: <https://ssrn.com/abstract=2855843>*, 2016.
- [74] K. Jerath, S. Netessine, and S. Veeraraghavan. Revenue management with strategic customers: Last-minute selling and opaque selling. In *Management Science*, volume 56, pages 430–448, 2010.
- [75] Y. Jiang. Price discrimination with opaque products. In *Journal of Revenue & Pricing Management*, volume 6, pages 118–134, 2007.
- [76] JuniperResearch. ‘loot boxes and skins gambling to generate a \$50 billion industry by 2022’. 2018. URL <https://www.juniperresearch.com/press/press-releases/loot-boxes-and-skins-gambling>.
- [77] K. C. Kiwiel. Convergence and efficiency of subgradient methods for quasiconvex minimization. *Mathematical programming*, 90(1):1–25, 2001.
- [78] Y. Liu and Z. J. Zhang. The benefits of personalized pricing in a channel. *Marketing Science*, 25(1):97–105, 2006.

- [79] W. Ma and D. Simchi-Levi. Reaping the benefits of bundling under high production costs. 2016.
- [80] K. S. Moorthy. Market segmentation, self-selection, and product line design. *Marketing Science*, 3(4):288–307, 1984.
- [81] A. Muoz Medina and S. Vassilvitskii. Revenue optimization with approximate bid predictions. 06 2017.
- [82] C. Narasimhan. A price discrimination theory of coupons. *Marketing Science*, 3(2):128–147, 1984.
- [83] A. Obama. Big data and differential pricing. *White House Report*, 2016. URL https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf/.
- [84] Ö. Özer and Y. Zheng. Markdown or everyday low price? the role of behavioral motives. *Management Science*, 62(2):326–346, 2015.
- [85] M. Pai and R. Vohra. Optimal dynamic auctions. 2008.
- [86] G. Pavlov. Optimal mechanism for selling two goods. *The BE Journal of Theoretical Economics*, 11(1), 2011.
- [87] R. Phillips. Optimizing prices for consumer credit. *Journal of Revenue and Pricing Management*, 12(4):360–377, 2013.
- [88] I. Popescu. A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Math. Oper. Res.*, 30(3):632–657, Aug. 2005. ISSN 0364-765X. doi: 10.1287/moor.1040.0137. URL <http://dx.doi.org/10.1287/moor.1040.0137>.
- [89] D. Post. Variable opaque products in the airline industry: A tool to fill the gaps and increase revenues. *Journal of Revenue and Pricing Management*, 9(4):292–299, 2010.
- [90] D. Post and M. Spann. Improving airline revenues with variable opaque products: “blind booking” at germanwings. In *Interfaces*, volume 42, pages 329–338, 2012.
- [91] S. Qiang and M. Bayati. Dynamic pricing with demand covariates. *Available at SSRN: <https://ssrn.com/abstract=2765257>*, 2016.
- [92] J. Robinson. The economics of imperfect competition. *Journal of Political Economy*, 42(2):249–259, 1934.
- [93] S. C. Salop. Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141–156, 1979. ISSN 0361915X. URL <http://www.jstor.org/stable/3003323>.

- [94] R. Schmalensee. Output and welfare implications of monopolistic third-degree price discrimination. *The American Economic Review*, 71(1):242–247, 1981.
- [95] A. Shapiro. On duality theory of conic linear problems. In *Semi-infinite programming*, pages 135–165. Springer, 2001.
- [96] D. Shapiro and X. Shi. Market segmentation: The role of opaque travel agencies. *Journal of Economics & Management Strategy*, 17(4):803–837, 2008.
- [97] J.-j. Shih, C.-c. Mai, and J.-c. Liu. A general analysis of the output effect under third-degree price discrimination. *The Economic Journal*, 98(389):149–158, 1988.
- [98] X. Su. Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5):726–741, 2007.
- [99] O. Tamuz. A lower bound on seller revenue in single buyer monopoly auctions. *Operations Research Letters*, 41(5):474–476, 2013.
- [100] P. Tassi. China Has Forced Blizzard To Reveal Exact 'Overwatch' And 'Hearthstone' Drop Rates., 2017. <https://www.forbes.com/sites/insertcoin/2017/05/05/china-has-forced-blizzard-to-reveal-drop-rates/#615b69035a14>.
- [101] B. Tuttle. Flight prices to get personal? airfares could vary depending on who is traveling. *Time*, Mar 2013. URL <http://business.time.com/2013/03/05/flight-prices-to-get-personal-airfares-could-vary/>.
- [102] H. R. Varian. Price discrimination and social welfare. *The American Economic Review*, 75(4):870–875, 1985.
- [103] Y. Xiao and J. Chen. Evaluating the potential effects from probabilistic selling of similar products. *Naval Research Logistics (NRL)*, 61(8):604–620, 2014.
- [104] X. K. Xie, R. Verma, and C. K. Anderson. Demand growth in services: A discrete choice analysis of customer preferences and online selling. *Decision Sciences*, 47(3):473–491, 2016.
- [105] X. Xu, Z. Lian, X. Li, and P. Guo. A hotelling queue model with probabilistic service. *Operations Research Letters*, 44(5):592–597, 2016.
- [106] Z. Xu and A. J. Dukes. Price discrimination in a market with uninformed consumer preferences. Available at SSRN: <https://ssrn.com/abstract=2777081>, 2016.
- [107] D. Zendle and P. Cairns. Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PloS one*, 13(11):e0206767, 2018.
- [108] J. Zhang. The perils of behavior-based personalization. *Marketing Science*, 30(1):170–186, 2011.

Appendices

Chapter A

Appendix to Chapter 1

A.1 A Primer on Discriminatory Pricing

In this work, we consider pricing strategies for customers with valuations drawn from an exchangeable distribution. Due to the symmetry of exchangeable valuation distributions, it is natural to assume that the optimal pricing strategy would be to offer an identical price for each item. Surprisingly, this is not the case; discriminatory pricing where some items are priced higher than others can yield significantly more revenue. (Note any permutation of the prices is also optimal.) Consider the following simple example where $N = 2$, and customers draw their item valuations from an i.i.d. distributions V_1 and V_2 :

$$V_1, V_2 \sim \begin{cases} 2 & : \text{w.p. } 1/3 \\ 1 & : \text{w.p. } 2/3 \end{cases}$$

The optimal single price strategy offers both items at price 2, and earns revenue of $\mathcal{R}_{SP} = 2 \cdot \mathbb{P}(\max\{V_1, V_2\} \geq 2) = \frac{10}{9}$. Now consider a discriminatory pricing where item 1 is sold at a price of 1 and item 2 is sold at a price of 2, then $\mathcal{R}_{DP}(1, 2) = 1 \cdot (\mathbb{P}(V_1 = 1, V_2 = 1) + \mathbb{P}(V_1 = 2, V_2 = 1)) + \mathbb{P}(V_1 = 2, V_2 = 2) + 2 \cdot \mathbb{P}(V_1 = 1, V_2 = 2) = \frac{11}{9}$ (note that the higher price item is purchased in the event of a tie w.l.o.g.). This extra $\frac{1}{9}$ is from two opposing forces at play. The low price of 1 allows the discriminatory strategy to extract revenue from customers with low valuations for both items, an additional expected revenue of $\frac{4}{9}$. The downside of the

low price is that it cannibalizes sales from the high priced item when $V_1 = 2$, resulting in a loss of $\frac{1}{3}$. Overall, the upside outweighs the downside and increased revenue can be had by offering different prices for these i.i.d. valued items.

Given that the revenue from discriminatory pricing can exceed single pricing, a natural question is then how much more can revenue can discriminatory pricing earn. In i.i.d. settings, the question has been fully resolved.

Lemma A.1.1 (Dutting and Klimm [49] Theorems 3 and 4). *Let N be the number of items. When valuations are drawn i.i.d.,*

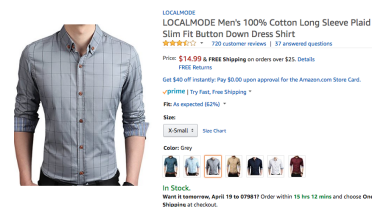
$$\mathcal{R}_{DP} \leq \left(2 - \frac{1}{N}\right) \mathcal{R}_{SP}.$$

Further, this bound is tight for each N .

In addition to being theoretically interesting, discriminatory pricing strategies are common in online marketplaces, even for basic retail goods. In Fig. A.1, two nearly identical shirts are offered for different prices on Amazon.com.



(a) Dark blue shirt for \$17.88.



(b) Light blue shirt for \$14.99.

Figure A.1

A.2 An Extension of Exchangeability

Here we describe Salop's circle model, a core model for horizontally differentiated items, and discuss how Theorem 1.3.1(i) can be extended to this case. In Salop's circle model, each item is represented as a point on a circle. The N points are equidistant on

the circumference of the circle, and denoted by y_1, \dots, y_N . Each customer corresponds to a random point X on the circumference of the circle, distributed uniformly at random. The customer then values each item i according to $V_i = a - b\|X - y_i\|$, where the norm corresponds the distance traveled on the circle and a, b are tunable parameters. Note that the underlying joint distribution for item valuations arising from Salop's circle is not independent, as closeness to one item on the circle necessarily implies the customer is farther from the other items.

When $N = 2$, Salop's circle reduces to the Hotelling model, which is well-known to be exchangeable. We now show that when $N = 3$, Salop's circle still gives rises to an exchangeable valuation distribution. First, observe from Fig. A.2 that the six possible valuation orders of the 3 items are equally likely. For example, if a customer is in region c, they prefer item 3, then 1, then item 2. If a customer is in region d, they prefer item 3, then item 2, then item 1. All 6 orderings are possible when $N = 3$. Since X is drawn uniformly at random, then exchangeability follows immediately.

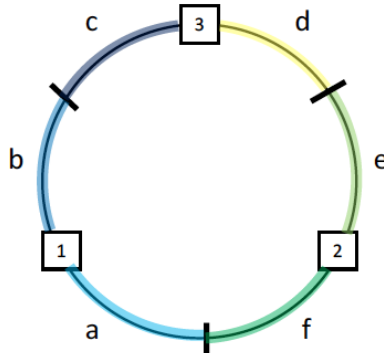


Figure A.2: Salop's circle for $N=3$ products, divided into six equal regions.

Unfortunately this argument for exchangeability does not extend beyond $N = 3$. One easy way to see this is to note that, when $N = 4$, the customers are partitioned into 8 regions corresponding to 8 possible valuation orderings (analogous to Figure A.2). However, there are 24 possible valuation orderings when $N = 4$, and thus it cannot be the case that every permutation yields identical marginal distributions.

More generally, Salop's circle has $2N$ possible orderings arising from $2N$ customer regions, while exchangeability requires all $N!$ valuation orderings are equally likely.

To capture Salop's circle, we relax our definition of exchangeability from requiring every permutation to yield identical joint distributions to only a subset of permutations, S , to have identical joint distributions.

Definition A.2.1. *Let Σ be the set of all permutations on $\{1, \dots, N\}$. Let $\text{sort}(\cdot)$ be the function that sorts a vector in descending order. We call the random valuation vector $V = (V_1, \dots, V_N)$ S -exchangeable if S is a subset Σ such that for all $\sigma \in S$, $\sigma(\text{sort}(V))$ has the same joint distribution as V . \square*

In essence, S -exchangeability limits the possible orderings of the valuations for items to the set S and enforces that inside S , each of those orderings is equally likely. Note that when $S = \Sigma$, then S -exchangeability is equivalent to our earlier notion of exchangeability. It is now easy to see that Salop's circle model is S -exchangeable, where S describes the orderings arising from the $2N$ regions. The joint distributions for each of these orderings are all the same since X is uniformly distributed on the circle.

Finally, we note that when customers are pessimistic and their valuations are S -exchangeable, opaque selling always dominates the revenue from discriminatory pricing. This implies Corollary A.2.1, which is an extension of Theorem 1.3.1(i).

Corollary A.2.1. *Assume $\alpha = 1$ and customers valuations are S -exchangeable, then*

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}.$$

Proof. The proof is exactly the same as Theorem 1.3.1(i) where Σ , the set of all permutations, is replaced by S instead. \square \square

A.3 Missing Examples from Chapter 1

Example A.3.1 (OPQ Earning Order N Times More Revenue than SP). We construct an exchangeable distribution F over N items such that $\mathcal{R}_{SP} \leq 2$ and $\mathcal{R}_{OPQ} \geq \alpha N$, implying a gap between OPQ and SP on the order of αN . Note by Theorem 1.3.3(i) this implies $\mathcal{R}_{DP} \leq 2(1 + \log(N))$, and thus $\mathcal{R}_{OPQ} \geq \frac{\alpha}{2} \frac{N}{1 + \log(N)} \mathcal{R}_{DP}$.

To construct $V \sim F$, we specify $N + 1$ possible valuation vectors in \mathbb{R}^N and assume that each permutation of the specified valuation vector is equally likely. The first vector has valuations where one item is valued at 2^N and all others are valued at 0. The second vector has valuations where two items are valued at 2^{N-1} and all other are zero and so on. All vectors have probabilities chosen so that $\mathcal{R}_{SP}(2^{N-i}) \leq 2$. Formally,

$$V = \begin{cases} \text{Uniformly some permutation of } (2^N, 0, \dots, 0) & : w.p. \ 2^{-N} \\ \text{Uniformly some permutation of } (2^{N-1}, 2^{N-1}, 0, \dots, 0) & : w.p. \ 2^{-(N-1)} \\ \dots & \\ \text{Uniformly some permutation of } (2, 2, \dots, 2, 2) & : w.p. \ 2^{-1} \\ \text{Uniformly some permutation of } (1, 1, \dots, 1, 1) & : w.p. \ 2^{-N} \end{cases}$$

Then $\mathcal{R}_{SP} = \max_i 2^{N-i} \mathbb{P}(V^{(1)} \geq 2^{N-i}) = \max_i 2^{N-i} \sum_{j=N-i}^N 2^{-j} \leq \max_i 2^{N-i} 2^{-N+i+1} = 2$. Now we show $\mathcal{R}_{OPQ} \geq \alpha N$. Let prices be $p^i = 2^{N-i+1}$. Now consider a pessimistic customer with $V^{(1)} = V^{(2)} = \dots = V^{(i)} = 2^{N-i+1}$, which occurs w.p. $\alpha 2^{-(N-i+1)}$. By construction $V^{(i+1)} = 0$ and thus this customer purchases the size i opaque product at price 2^{N-i+1} . The total revenue is then

$$\mathcal{R}_{OPQ}(2^N, 2^{N-1}, \dots, 2) \geq \alpha \sum_{i=1}^N 2^i 2^{-i} = \alpha N.$$

Thus OPQ earns at least αN from pessimistic customers under this pricing strategy

and $\mathcal{R}_{OPQ} \geq \alpha N$. \square

Example A.3.2 (Tightness of Theorem 1.4.2 when $\alpha = 1$). *We describe a distribution F such that $\mathcal{R}_{1OPQ}^F/\mathcal{R}_{SP}^F = 2$ when customers are pessimistic, demonstrating tightness of Theorem 1.4.2. Fix $z \in (0, 1)$ and let V_1, \dots, V_N be i.i.d. where*

$$V_i = \begin{cases} 1 & : w.p. \ z \\ 1 - (1 - z)^N & : w.p. \ 1 - z \end{cases}$$

Then,

$$\mathcal{R}_{SP} = \mathcal{R}_{SP}(1) = \mathcal{R}_{SP}(1 - (1 - z)^N) = 1 - (1 - z)^N.$$

Similarly we compute the revenue of a 1OPQ strategy with prices $(1, 1 - (1 - z)^N)$. In this strategy, the opaque product is only purchased if the customer has a high valuation for all items or a low valuation for all items. Thus,

$$\mathcal{R}_{1OPQ}(1, 1 - (1 - z)^N) = 1(1 - z^N - (1 - z)^N) + (1 - (1 - z)^N)(z^N + (1 - z)^N).$$

Then,

$$\begin{aligned} \frac{\mathcal{R}_{1OPQ}}{\mathcal{R}_{SP}} &\geq \frac{\mathcal{R}_{1OPQ}(1, 1 - (1 - z)^N)}{\mathcal{R}_{SP}} = \frac{(1 - z^N - (1 - z)^N) + (1 - (1 - z)^N)(z^N + (1 - z)^N)}{1 - (1 - z)^N} \\ &= 1 + z^N + (1 - z)^N - \frac{z^N}{1 - (1 - z)^N}. \end{aligned} \quad (\text{A.1})$$

Then Eq. (A.1) can be arbitrarily close to 2 as z goes to zero. Note this example holds for any N . \square

Example A.3.3 (Tightness of Theorem 1.4.2 when $\alpha = 0$). *We describe a distribution F such that $\mathcal{R}_{1OPQ}^F/\mathcal{R}_{SP}^F = \frac{3}{2}$ when customers are risk-neutral, demonstrating tightness of Theorem 1.4.2 when $N = 2$. Let valuations for the two items be drawn*

i.i.d. according to the CDF

$$\mathbb{P}(V_i \leq x) = \sqrt{1 - \frac{1}{x}}, \quad x \in [1, \infty).$$

Note that for any $x \in [1, \infty)$, we have that $\mathbb{P}(\max\{V_1, V_2\} \leq x) = 1 - \frac{1}{x}$ and $\mathbb{P}(\max\{V_1, V_2\} \geq x) = \frac{1}{x}$. Thus $R_{SP} = 1$ since $p\mathbb{P}(\max\{V_1, V_2\} \geq p) = 1 \forall p \in [1, \infty)$.

Now consider the 1OPQ strategy $(p, 1)$ for $p > 1$. To compute $\mathcal{R}_{1OPQ}(p, 1)$, let $V^{(1)} = \max_i\{V_i\}$, $V^{(2)} = \min_i\{V_i\}$, $u = V^{(1)} - p$, and $u^2 = \frac{V^{(1)} + V^{(2)}}{2} - 1$. Note that u and u^2 are the utilities of buying the best item or the opaque product, respectively. By conditioning on the event $V^{(1)} \geq p$, we show that

$$\begin{aligned} \frac{\mathcal{R}_{1OPQ}}{\mathcal{R}_{SP}} &\geq \mathcal{R}_{1OPQ}(p, 1) \\ &= \mathbb{P}(V^{(1)} \geq p) \left(p\mathbb{P}(u \geq (u^2)^+ | V^{(1)} \geq p) + \mathbb{P}(u^2 > (u)^+ | V^{(1)} \geq p) \right) \\ &\quad + \mathbb{P}(V^{(1)} < p) \left(p\mathbb{P}(u \geq (u^2)^+ | V^{(1)} < p) + \mathbb{P}(u^2 > (u)^+ | V^{(1)} < p) \right) \\ &= \mathbb{P}(V^{(1)} \geq p) (p\mathbb{P}(V^{(1)} - V^{(2)} \geq 2p - 2 | V^{(1)} \geq p) + \mathbb{P}(V^{(1)} - V^{(2)} < 2p - 2 | V^{(1)} \geq p)) \\ &\quad + \mathbb{P}(V^{(1)} < p) (p\mathbb{P}(V^{(1)} - V^{(2)} \geq 2p - 2 | V^{(1)} < p) + \mathbb{P}(V^{(1)} - V^{(2)} < 2p - 2 | V^{(1)} < p)) \\ &= \mathbb{P}(V^{(1)} \geq p) (p\mathbb{P}(V^{(1)} - V^{(2)} \geq 2p - 2 | V^{(1)} \geq p) + \mathbb{P}(V^{(1)} - V^{(2)} < 2p - 2 | V^{(1)} \geq p)) \\ &\quad + \mathbb{P}(V^{(1)} < p) \\ &= \frac{1}{p} (p\mathbb{P}(V^{(1)} - V^{(2)} \geq 2p - 2 | V^{(1)} \geq p) + \mathbb{P}(V^{(1)} - V^{(2)} \leq 2p - 2 | V^{(1)} \geq p)) + (1 - \frac{1}{p}) \\ &\geq \mathbb{P}(V^{(1)} \geq 2p - 2 | V^{(1)} \geq p) + \frac{1}{p} \mathbb{P}(V^{(1)} - V^{(2)} \leq 2p - 2 | V^{(1)} \geq p) + 1 - \frac{1}{p} \\ &= \frac{\frac{1}{2p-2}}{\frac{1}{p}} + \frac{1}{p} \mathbb{P}(V^{(1)} - V^{(2)} \leq 2p - 2 | V^{(1)} \geq p) + 1 - \frac{1}{p} \end{aligned} \tag{A.2}$$

The first inequality follows since $\mathcal{R}_{SP} = 1$ and $(p, 1)$ is feasible for 1OPQ. The first equality follows from the definition of $\mathcal{R}_{1OPQ}(p, 1)$ conditioning on $V^{(1)} \geq p$. The second equality follows from the definitions of u and u^2 . The third equality follows from the fact that $V^{(2)} \geq 1$ combined with the case where $V^{(1)} < p$. The fourth equality follows from the

distribution F . The second inequality follows since $V^{(2)} \geq 0$. The last equality follows from Bayes rule. As p goes to ∞ , the expression in Eq. (A.2) goes to $\frac{3}{2}$, matching the upper bound in Theorem 1.4.2. \square

A.4 Additional Proofs from Chapter 1

Proof of Corollary 1.3.1.

Case (i): From the proof of Theorem 1.3.1(i), the revenue from OPQ dominates the revenue under SP conditional on every ordering of the valuations. When $V_N = V^{(1)}$, only the highest-valued item can be purchased under DP. Conditional on this event, the revenue from DP is clearly at most \mathcal{R}_{SP} . Thus,

$$\mathcal{R}_{DP} \leq \frac{N-1}{N} \mathcal{R}_{OPQ} + \frac{1}{N} \mathcal{R}_{SP} = \frac{N-1}{N} \mathcal{R}_{OPQ} + \frac{1}{N+N\gamma} \mathcal{R}_{DP}.$$

Rearranging the inequality gives the result.

Case (ii): Recall from the proof Theorem 1.3.1(ii) that w.l.o.g. we may assume F is supported on $\{1, 1+\delta\}$. Suppose the OPQ prices are $p^i = 1+\delta$ for $i < N$ and $p^N = 1$. Recall that U is a r.v. denoting the number of items for which the customer has a valuation of $1+\delta$. A pessimistic customer buys an item at a price of $1+\delta$ if $1 \leq U \leq N-1$ and the opaque product of size N at price 1 otherwise. Letting $u_i = \mathbb{P}(U = i)$, then

$$\mathcal{R}_{OPQ} \geq 1 + \delta(1 - u_0 - u_N). \tag{A.3}$$

Similarly from Eq. (A.6) in Appendix A.4 we have that

$$\mathcal{R}_{DP} = 1 + \delta \sum_{i=1}^{N-1} \frac{N-i}{N} u_i \leq 1 + \delta \frac{N-1}{N} (1 - u_0 - u_N).$$

Now since $\mathcal{R}_{DP} = (1 + \gamma)\mathcal{R}_{SP}$ and $\mathcal{R}_{SP} \geq \mathcal{R}_{SP}(1) = 1$, then

$$\delta \frac{N-1}{N} (1 - u_0 - u_N) \geq \gamma \mathcal{R}_{SP}.$$

Combining these inequalities with Eq. (A.3) we obtain

$$\begin{aligned} \mathcal{R}_{OPQ} &\geq 1 + \delta(1 - u_0 - u_N) \\ &\geq \mathcal{R}_{DP} + \frac{1}{N} \delta(1 - u_0 - u_N) \\ &\geq \mathcal{R}_{DP} + \frac{\gamma}{N-1} \mathcal{R}_{SP} = \left(1 + \frac{\gamma}{(N-1)(1+\gamma)}\right) \mathcal{R}_{DP} \end{aligned}$$

When customers are risk-neutral, they always purchase the opaque product of size N garnering revenue 1. Thus $\mathcal{R}_{OPQ} \geq \alpha \left(1 + \frac{\gamma}{(N-1)(1+\gamma)}\right) \mathcal{R}_{DP} + 1 - \alpha$. Thus when $\alpha \geq 1 - \frac{\gamma}{\gamma N + N - 1}$ we obtain $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$. \square

Proof of Theorem 1.3.3

Proof. First we prove part (i). Fix an exchangeable distribution F over N items. For ease of exposition we assume F is continuous and ignore ties, although the same argument follows when F is not continuous and one carefully considers the tie-breaking procedure. Let $p_1 \geq p_2 \geq \dots \geq p_N$ be the optimal discriminatory pricing and let q_1, q_2, \dots, q_N be the probability item $1, 2, \dots, N$ is sold under this pricing. Define $Q_i = \{v \in \text{supp}(F) \mid v_i - p_i \geq (v_j - p_j)^+ \forall j\}$. Note that $q_i = \mathbb{P}(Q_i)$. Let $\sigma_{i,j} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ be the map that interchanges v_i and v_j in a vector v .

We first observe that $\mathbb{P}(\sigma_{i,j}(Q_i)) = q_i$ by exchangeability. Further, for all $i < j$ and any $v \in Q_i$, notice that $\sigma_{i,j}(v) \notin Q_i$ since $p_i \geq p_j$. With that established note,

$$\mathbb{P}(V^{(1)} \geq p_i) \geq \mathbb{P}(\cup_{j \geq i} \sigma_{i,j}(Q_i)) = \cup_{j \geq i} \mathbb{P}(\sigma_{i,j}(Q_i)) = (N - i + 1)q_i.$$

In words, the above equation says the probability the highest valuation is greater

than p_i is lower bounded by the probability of selling item i under a discriminatory pricing, union with disjoint permutations of the event. Using this observation we can bound the gap between \mathcal{R}_{SP} and \mathcal{R}_{DP} as

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{SP}} \leq \frac{\sum_{i=1}^N q_i p_i}{\max_i \mathbb{P}(V^{(1)} \geq p_i) p_i} \leq \frac{\sum_{i=1}^N q_i p_i}{\max_i (N - i + 1) q_i p_i} \leq H_N.$$

The final inequality follows from Lemma A.5.1 in Appendix A.5 with $C_1 = 0, C_2 = 1$, $K = 1$, and H_N is the N^{th} harmonic number. Recalling the fact that $H_N \leq 1 + \log(N)$ yields the result.

Part (ii) follows from the observation that one of the N prices in an optimal opaque selling strategy garners the most revenue. Call that price p_i and let q_i be the probability an opaque product of size i is sold. Then we must have that $q_i \leq \mathbb{P}(V^{(1)} \geq p_i)$ and thus $p_i q_i \leq \mathcal{R}_{SP}(p_i) \leq \mathcal{R}_{SP}$. Therefore, $\mathcal{R}_{OPQ} \leq N p_i q_i \leq N \mathcal{R}_{SP}$.

Part (iii) follows from combining Example A.3.1 with part (i). \square

Proof of Lemma 1.3.3

The proof is based on following lemma which makes a fundamental connection between lottery pricings and the Myerson auctions.

Lemma A.4.1 (Lemmas 3 and 4 in Chawla et al. [32]). *Consider a customer with a valuation draw \vec{v} and let $i^* = \operatorname{argmax}_i v_i$. Let \mathcal{L} be a lottery pricing such that the customer buys lottery $(p, q_1, \dots, q_N) = l \in \mathcal{L}$. Let \mathcal{M} be the Myerson auction for one item, run on N bidders with valuations drawn from F . Then*

$$\mathcal{R}_{\mathcal{L}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \sum_{i \neq i^*} q_i v_i.$$

where $\mathcal{R}_{\mathcal{L}}(\vec{v})$ and $\mathcal{R}^{\mathcal{M}}(\vec{v})$ denote the revenue earned by the lottery pricing and Myerson auction when the valuation draw is \vec{v} .

Proof. Let (p, \vec{p}) denote the prices of an optimal OPQ strategy under F , where p is the price of items and \vec{p} are the prices of the opaque products. Note that every opaque product S can be written as a lottery with the same price and a uniform allocation probability over S . Furthermore, we can describe the items as N individual lotteries, each priced at p with a deterministic allocation. Thus for risk-neutral customers, our opaque selling strategy can be recast as a lottery pricing which we call \mathcal{L}_{OPQ} , i.e.,

$$\mathcal{R}_{OPQ} = \mathcal{R}_{\mathcal{L}_{OPQ}}.$$

From Lemma A.4.1, we have that

$$\mathcal{R}_{\mathcal{L}_{OPQ}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \sum_{i \neq i^*} q_i v_i. \quad (\text{A.4})$$

We note that if a customer with valuation \vec{v} and $i^* = \operatorname{argmax}_i v_i$ purchases an item, then $\sum_{i \neq i^*} q_i v_i = 0$ since $q_{i^*} = 1$. Otherwise if an opaque product is purchased, $\sum_{i \neq i^*} q_i \leq \frac{N-1}{N}$ and $v_i \leq v^{(2)}$ for $i \neq i^*$. Combining these facts with Eq. (A.4) yields

$$\mathcal{R}_{\mathcal{L}_{OPQ}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \frac{N-1}{N} v^{(2)}. \quad (\text{A.5})$$

Thus

$$\begin{aligned} \mathcal{R}_{OPQ} &\leq \mathcal{R}^{\mathcal{M}} + \frac{N-1}{N} E[V^{(2)}] \\ &\leq \mathcal{R}^{\mathcal{M}} + \frac{N-1}{N} \mathcal{R}^{\mathcal{M}} \\ &\leq (4 - \frac{2}{N}) \mathcal{R}_{SP} \end{aligned}$$

The first inequality follows from taking the expectation of Eq. (A.5) over \vec{v} . The second inequality follows from the fact that $E[V^{(2)}]$ is the revenue of a second price auction, which is dominated by the Myerson auction. The third inequality follows

from Lemma 1.3.2. \square

\square

Proof of Theorem 1.4.1

Proof. Let F be an exchangeable distribution where item valuations can take only two points $\{a, b\}$ where $a < b$, and suppose the market is α -mixed. Recall for distributions supported on two points the optimal discriminatory pricing uses prices $\vec{p} = (a, a, \dots, a), (b, b, \dots, b)$ or a mixed pricing where exactly one price (since F is exchangeable it doesn't matter which price) is low (a, b, b, \dots, b) . If either (a, a, \dots, a) or (b, b, \dots, b) is the the optimal discriminatory pricing given F , then $\mathcal{R}_{SP} = \mathcal{R}_{DP}$ and the claim follows automatically. Suppose $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, then the optimal pricing is the mixed strategy and, under a mixed pricing, a discriminatory selling strategy always sells the item. Further we restrict ourselves to 1OPQ strategies that always sell the item, thus we may normalize the support of F to $\{1, 1 + \delta\}$ without changing the ratio $\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}}$.

Define U to be the random variable supported on $\{1, \dots, N\}$ such that $\mathbb{P}(U = i) = \mathbb{P}(V^{(i)} = 1 + \delta, V^{(i+1)} = 1)$ where $V^{(i)}$ is the i^{th} highest order statistic of F . In words, U is the random variable for how many of the N valuations are equal to $1 + \delta$. Recall since F is exchangeable, if $U = i$ then all $\binom{N}{i}$ arrangements of valuations over the N items are equally likely (thus knowing the distribution of U is equivalent to knowing F in a two point setting). Let $u_i := \mathbb{P}(U = i)$. Conditioning on U we can compute \mathcal{R}_{DP} as

$$\begin{aligned} \mathcal{R}_{DP} &= E[\mathcal{R}_{DP}(1, 1 + \delta, \dots, 1 + \delta) | U = i] \\ &= u_0 + \sum_{i=1}^N u_i \left(\frac{i}{N} + \frac{N-i}{N} (1 + \delta) \right) = 1 + \delta \sum_{i=1}^N \frac{N-i}{N} u_i \end{aligned} \quad (\text{A.6})$$

where the second equality follows from Eq. (1.1).

For each $i = 1, \dots, N$, we shall lower bound the 1OPQ pricing strategies $(p, p_N) =$

$(1 + \frac{N-i}{N}\delta, 1)$. To analyze the revenue from such a pricing, first note a pessimistic customers always buy an item at price $1 + \frac{N-i}{N}\delta$ for any 1OPQ pricing $(1 + \frac{N-i}{N}\delta, 1)$ as long as $U \neq 0, N$, and the opaque product otherwise. Thus a pessimistic customer has expected revenue $1 + (1 - u_0 - u_N)\frac{N-i}{N}\delta$.

For risk-neutral customers and 1OPQ pricing $(1 + \frac{N-i}{N}\delta, 1)$, customers will purchase an item at price $1 + \frac{N-i}{N}\delta$ if $0 < U \leq i, U \neq N$, and the opaque product otherwise. Thus a risk-neutral customer has expected revenue $1 + \frac{N-i}{N}\delta \sum_{j=1}^{\min(N-1, i)} u_j$. Putting them together we have

$$\begin{aligned}
\mathcal{R}_{1OPQ} &\geq 1 + \max_{i \in \mathcal{N}} \alpha (1 - u_0 - u_N) \frac{N-i}{N} \delta + (1 - \alpha) \frac{N-i}{N} \delta \sum_{j=1}^{\min(N-1, i)} u_j \\
&= 1 + \max_{i \in \mathcal{N}} \frac{N-i}{N} \delta (1 - u_0 - u_N) \left(\alpha + \frac{1 - \alpha}{1 - u_0 - u_N} \sum_{j=1}^{\min(N-1, i)} u_j \right) \\
&\geq 1 + \max_{i \in \mathcal{N}} \frac{N-i}{N} \delta \sum_{j=1}^{\min(N-1, i)} u_j \\
&= 1 + \max_{i \in \mathcal{N}} \frac{N-i}{N} \delta \sum_{j=1}^i u_j \tag{A.7}
\end{aligned}$$

where the second inequality follows from noting Eq. (A.7) is an increasing function of α and then plugging in $\alpha = 0$. The second equality follows since $i = N$ is not the maximizer. Let $u'_i = \frac{u_i}{1 - u_0}$ so that $\sum_{i=1}^N u'_i = 1$. Then

$$\begin{aligned}
\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} &\leq \frac{1 + \delta \sum_{i=1}^N \frac{N-i}{N} u_i}{1 + \delta \max_{i \in \mathcal{N}} \frac{N-i}{N} \sum_{j=1}^i u_j} = \frac{1 + \delta(1 - u_0) \sum_{i=1}^N \frac{N-i}{N} u'_i}{1 + \delta(1 - u_0) \max_{i \in \mathcal{N}} \frac{N-i}{N} \sum_{j=1}^i u'_j} \\
&\leq \frac{1 + \frac{(1-u_0)\delta}{N} H_{N-1}}{1 + \frac{(1-u_0)\delta}{N}}.
\end{aligned}$$

The first inequality follows from Eq. (A.6) and Eq. (A.7). The second inequality follows from applying Lemma A.5.2 (under the appropriate change of variables i.e. relabeling $x_i \rightarrow x_{N-i}$) where H_{N-1} is the $(N-1)^{th}$ harmonic number. Lastly note

that

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} \leq \frac{\mathcal{R}_{DP}}{\mathcal{R}_{SP}(1+\delta)} \leq \frac{1+\delta(1-u_0)}{(1-u_0)(1+\delta)} \leq \frac{1+\delta(1-u_0)}{(1-u_0)\delta}.$$

The first inequality follows from observing that \mathcal{R}_{1OPQ} earns as much as a SP strategy with price $1+\delta$, which has expected revenue $(1-u_0)(1+\delta)$. The second inequality follows from a simple upper bound on Eq. (A.6). Define $C = (1-u_0)\delta$ and putting it all together we have,

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} \leq \max_{N \in \mathbb{N}, C > 0} \min \left(\frac{1+C}{C}, \frac{1+C \frac{1+H_{N-1}}{N}}{1+\frac{C}{N}} \right) \quad (\text{A.8})$$

which can be checked to be maximized when $N = 7, C = \frac{2}{29} (5 + 4\sqrt{65})$ yielding a ratio $\geq .719$. \square

Proof of Theorem 1.4.2.

Proof. We divide the proof of Theorem 1.3.2 into two parts. First we show that when customers are purely pessimistic, $\mathcal{R}_{1OPQ} \leq 2\mathcal{R}_{SP}$. Second we show that when customers are purely risk-neutral $\mathcal{R}_{1OPQ} \leq (2 - \frac{1}{N}) \mathcal{R}_{SP}$. To obtain the result, we relax 1OPQ to observe X_α and price pessimistic and risk-neutral customers separately. Using the previously mentioned results, we get that $\mathcal{R}_{1OPQ} \leq (\alpha \cdot 2 + (1-\alpha)(2 - \frac{1}{N})) \mathcal{R}_{SP}$ which is equivalent to the desired result.

First we prove that $\mathcal{R}_{1OPQ} \leq 2\mathcal{R}_{SP}$ when customers are all pessimistic ($\alpha = 1$). Fix a distribution F and let (p, p^N) be an optimal solution corresponding to \mathcal{R}_{1OPQ}^F .

Then

$$\begin{aligned}
\mathcal{R}_{1OPQ} &= \mathcal{R}_{1OPQ}(p, p^N) \\
&= p\mathbb{P}(\max_i\{V_i - p\} \geq \min_i\{V_i - p^N, 0\}) \\
&\quad + p^N\mathbb{P}(V^N - p^N > \max_i\{V_i - p\} \cap V^N - p^N \geq 0) \\
&\leq p\mathbb{P}(\max_i\{V_i - p\} \geq 0) + p^N\mathbb{P}(\max_i\{V_i - p^N\} \geq 0) \\
&= \mathcal{R}_{SP}(p) + \mathcal{R}_{SP}(p^N) \\
&\leq 2\mathcal{R}_{SP}.
\end{aligned}$$

The second equation follows from the definition of $\mathcal{R}_{1OPQ}(p, p^N)$ and breaks ties by choosing to buy an item versus an opaque product. The first inequality follows from increasing the size of the event being measured. The second inequality follows from the fact that p and p^N are feasible solutions to SP. For tightness, see Example A.3.2.

Now we focus on the case when customers are all risk-neutral ($\alpha = 0$) and show $\mathcal{R}_{1OPQ} \leq (2 - \frac{1}{N})\mathcal{R}_{SP}$. Fix a distribution F and let (p, p^N) be an optimal solution corresponding to \mathcal{R}_{1OPQ}^F . Our proof breaks into two cases depending on the relative gap between p and p^N , corresponding to $p^N \geq \frac{1}{N}p$ (Case 1) and $p^N < \frac{1}{N}p$ (Case 2). Fig. A.3 illustrates the geometric difference in the two cases for $N = 2$ items.

Case 1: Recall in this case, $p^N \geq \frac{1}{N}p$. We first define q_A, q_B, q_C, q_D to be the probabilities corresponding to the following disjoint events under F , namely

$$\begin{aligned}
q_A &= \mathbb{P}(\max\{V_i\} - p \geq \frac{\sum V_i}{N} - p^N, \max\{V_i\} \geq p) \\
q_B &= \mathbb{P}(\max\{V_i\} - p < \frac{\sum V_i}{N} - p^N, \max\{V_i\} \geq p) \\
q_C &= \mathbb{P}(\max\{V_i\} < p, \frac{\sum V_i}{N} \geq p^N) \\
q_D &= \mathbb{P}(\max\{V_i\} < p, \frac{\sum_{i=1}^N V_i}{N} < p^N).
\end{aligned}$$

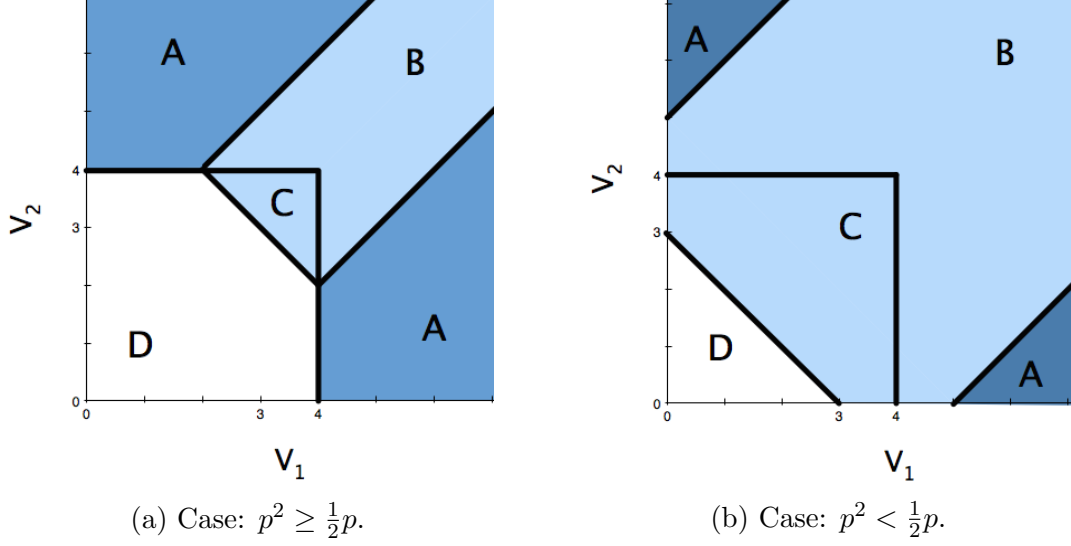


Figure A.3: These figures demonstrate the partition of the valuation space for a risk-neutral customer when $N = 2$. In both figures, $p = 4$. The left figure corresponds to a small discount for the opaque product and the right figure corresponds to a large discount for the opaque product. The four letters denote different buying behaviors of the customer under a (p, p^N) 1OPQ strategy and a SP strategy with price p .

Note that $q_A + q_B + q_C + q_D = 1$. Using these probabilities, we have that $\mathcal{R}_{1OPQ}^F(p, p^N) = pq_A + p^N(q_B + q_C)$. Further, we can express the revenues from the single pricing approximately as $\mathcal{R}_{SP}^F(p) = p(q_A + q_B)$ and $\mathcal{R}_{SP}^F(p^N) = p^N \mathbb{P}(\max\{V_i\} \geq p^N) \geq p^N(q_A + q_B + q_C)$. Thus we have

$$\frac{\mathcal{R}_{1OPQ}^F}{\mathcal{R}_{SP}^F} \leq \frac{\mathcal{R}_{1OPQ}^F(p, p^N)}{\max\{\mathcal{R}_{SP}^F(p), \mathcal{R}_{SP}^F(p^N)\}} \quad (\text{A.9})$$

$$\leq \frac{pq_A + p^N(q_B + q_C)}{\max\{p(q_A + q_B), p^N(q_A + q_B + q_C)\}} \quad (\text{A.10})$$

$$\leq \max_{\substack{a, b, c, d \geq 0 \\ a+b+c+d=1 \\ x \geq y \geq \frac{x}{N} \geq 0}} \frac{xa + y(b + c)}{\max\{x(a + b), y(a + b + c)\}}. \quad (\text{A.11})$$

The first inequality follows from the fact that p and p^N are feasible for SP. The second inequality follows from the previous discussion. The third inequality follows from the fact that $(p, p^N, q_A, q_B, q_C, q_D)$ is a feasible solution to the optimization problem in Eq. (A.11), which we denote by OPT . Lemma A.5.3, proved separately,

shows that $OPT \leq 2 - \frac{1}{N}$. Combining Lemma A.5.3 with Equations Eq. (A.9)-Eq. (A.11) completes the proof for the case of $p^N \geq \frac{1}{N}p$.

Case 2: Recall in this case, $p^N < \frac{1}{N}p$, where (p, p^N) are optimal prices corresponding to \mathcal{R}_{1OPQ}^F . We partition the valuation space under F according to the events

$$\begin{aligned} E_0 &= \{\max\{V_i\} < p^N\} \\ E_1 &= \{p^N \leq \max\{V_i\} < p\} \\ E_2 &= \{p \leq \max\{V_i\} < \frac{N}{N-1}(p - p^N)\} \\ E_3 &= \{\frac{N}{N-1}(p - p^N) \leq \max\{V_i\}\}. \end{aligned}$$

We upper bound the revenue from single opaque selling using this partition. Customers lying in E_0 do not generate any revenue. The revenue generated by customers lying in E_1 is at most $p^N \mathbb{P}(E_1)$ since they never consider buying an item at price p . The revenue generated by customers lying in E_3 is at most $p \mathbb{P}(E_3)$ since the best case scenario is that they all buy an item at price p . Lemma A.5.4, proved separately, shows that the customers lying in E_2 buy the opaque product at price p^N . Combining the previous arguments shows that

$$\mathcal{R}_{1OPQ}(p, p^N) \leq p \mathbb{P}(E_3) + p^N (\mathbb{P}(E_2) + \mathbb{P}(E_1)). \quad (\text{A.12})$$

Now suppose for contradiction that $\mathcal{R}_{1OPQ}(p, p^N) = \mathcal{R}_{1OPQ} > (2 - \frac{1}{N})\mathcal{R}_{SP}$. Then the following two inequalities must also hold:

$$\mathcal{R}_{1OPQ}(p, p^N) > (2 - \frac{1}{N})\mathcal{R}_{SP} \left(\frac{N}{N-1}(p - p^N) \right) = (2 - \frac{1}{N}) \left(\frac{N}{N-1}(p - p^N) \mathbb{P}(E_3) \right) \quad (\text{A.13})$$

$$\mathcal{R}_{1OPQ}(p, p^N) > (2 - \frac{1}{N})\mathcal{R}_{SP}(p^N) = (2 - \frac{1}{N})(p^N)(1 - \mathbb{P}(E_0)) \quad (\text{A.14})$$

by the optimality of \mathcal{R}_{SP} . Now define $\delta' = (p - p^N)/p$. Then combining the Eq. (A.13) and Eq. (A.14) with Eq. (A.12) and dividing by p yields

$$1 - \mathbb{P}(E_0) - \delta'(1 - \mathbb{P}(E_3) - \mathbb{P}(E_0)) = 1 - \delta' - (1 - \delta')\mathbb{P}(E_0) + \delta'\mathbb{P}(E_3) > \frac{2N-1}{N-1}\delta'\mathbb{P}(E_3) \quad (\text{A.15})$$

$$\begin{aligned} 1 - \mathbb{P}(E_0) - \delta'(1 - \mathbb{P}(E_3) - \mathbb{P}(E_0)) &= 1 - \delta' - (1 - \delta')\mathbb{P}(E_0) + \delta'\mathbb{P}(E_3) \\ &> \frac{2N-1}{N}(1 - \delta')(1 - \mathbb{P}(E_0)) \end{aligned} \quad (\text{A.16})$$

Rearranging Eq. (A.15) yields

$$1 - \delta' - (1 - \delta')\mathbb{P}(E_0) > \frac{N}{N-1}\delta'\mathbb{P}(E_3)$$

and rearranging Eq. (A.16) yields

$$1 - \delta' - (1 - \delta')\mathbb{P}(E_0) < \frac{N}{N-1}\delta'\mathbb{P}(E_3),$$

which is a contradiction and thus $\mathcal{R}_{1OPQ} \leq \frac{2N-1}{N}\mathcal{R}_{SP}$. For tightness when $N = 2$, see Example A.3.3. \square

Proof of Theorem 1.5.1.

Proof. The proof will depend on the following structural lemma which asserts that the prices can be found by carefully combing through the support of the valuation distribution. Suppose \vec{p} is the optimal price vector. By Lemma A.5.5, and for every $i \in \mathcal{N}$, either there exists a type j such that $p_i = v_{j,i} - \max_{k>i}\{(v_{j,k} - p_k)^+\}$, or no customer type buys item i . Using this observation, we can inductively enumerate the

prices starting from the lowest price and working upwards.

Algorithm 1: Enumerative Algorithm	
Main Enumerate Price Tree(F):	
	<i>Input:</i> Distribution F , supported on m types $v_i \in \mathbb{R}^N$.
	<i>Initialize:</i> $P^N = \cup_{j=1}^m v_{j,N}$
	for ($i = N - 1 : 1$) {
	for ($\tilde{p} \in P^{i+1}$) {
	$P^i = P^i \cup_{j=1}^m v_{j,i} - \max_{k>i} (v_{j,k} - \tilde{p}_k)^+$
	}
	}
	return $\arg \max_{\vec{p} \in P^1} \mathcal{R}_{DP}^F(\vec{p})$

We focus on optimal pricing for DP, and the same analysis holds for OPQ. Consider the following algorithm that proceeds by guessing the prices in order from low to high. Fix some ordering on the prices $p_1 \geq p_2 \geq \dots \geq p_N$, by exchangeability this is w.l.o.g. By Lemma A.5.5, it must be the case that the lowest price $p_N \in \{v_{j,N}\}_{j=1}^m$ or else that item N is not purchased by any customer. If that item is supposed to not be bought, we can set the price to ∞ effectively discarding the item. Thus there are $m + 1$ choices for p_N , one for each customer type and the ∞ no-purchase option. Under each of these choices, compute $\{\tilde{v}_{j,N-1}\}_{j=1}^m$ where $\tilde{v}_{j,i} = v_{j,i} - \max_{k>i} (v_{j,k} - p_k)^+$. Again, it must be the case that $p_{N-1} \in \{\tilde{v}_{j,N-1}\}_{j=1}^m$ by Lemma A.5.5 or else it is not bought and we can set the price to ∞ . Proceeding in this way we create a tree of size of depth N with $m + 1$ branches, terminating in $(m + 1)^N$ leaf nodes each corresponding to a potential optimal solution. For each leaf node, one can compute the revenue from the candidate price in linear time. Thus, the overall runtime is simply $\Theta(m^N)$. \square

A.5 Auxiliary Lemmas

Lemma A.5.1. *For any $x_1, \dots, x_N \geq 0$, $\sum_i x_i \leq K$, and constants $C_1, C_2 \geq 0$,*

$$\frac{C_1 + C_2 \sum_i x_i}{C_1 + C_2 \max_i ix_i} \leq \frac{C_1 + C_2 K H_N}{C_1 + C_2 K}. \quad (\text{A.17})$$

Proof. We first claim that the left hand side of Eq. (A.17), viewed as an optimization problem over all feasible \vec{x} is maximized when $ix_i = jx_j$ for all i, j . To prove this, suppose $\vec{x} \in [0, K]^N$ maximizes the left hand side of Eq. (A.17) and suppose the claim does not hold. Let $i = \arg \max_k ix_k$ and $j = \arg \min_k ix_k$, so by assumption $ix_i > jx_j$. Define y_i, y_j as solutions to the following system of two linear equations:

$$y_i + y_j = x_i + x_j$$

$$iy_i = jy_j$$

This yields $iy_i = jy_j = \frac{ij(x_i + x_j)}{i+j}$ which can be rewritten as $iy_i = jy_j = \frac{i}{i+j}jx_j + \frac{j}{i+j}ix_i$, a weighted average of ix_i and jx_j . Consider the maximal solution with components x_i, x_j replaced by y_i, y_j . Since $y_i + y_j = x_i + x_j$, the numerator in the l.h.s. of Eq. (A.17) is unchanged. However since $\max\{iy_i, jy_j\} < ix_i$, the denominator strictly decreases (in the case of many indices's that maximize ix_k , iterating the argument at most $N-1$ times yields a strict reduction) contradicting the optimality of \vec{x} . Thus Eq. (A.17) is maximized when $ix_i = jx_j \forall i, j$. Solving for the worst case \vec{x} gives $x_i = \frac{x_1}{i}$, and plugging in gives

$$\max_{\vec{x}} \frac{C_1 + C_2 \sum_i x_i}{C_1 + C_2 \max_i ix_i} \leq \frac{C_1 + C_2 x_1 \sum_{i=1}^N \frac{1}{i}}{C_1 + C_2 x_1} \leq \frac{C_1 + C_2 K H_N}{C_1 + C_2 K}. \quad (\text{A.18})$$

□

Lemma A.5.2. For any $x_1, \dots, x_{N-1} \geq 0$, such that $\sum_i x_i \leq 1$, and constant $C \geq 0$,

$$\frac{1 + C \sum_{i=1}^{N-1} i x_i}{1 + C \max_i i \sum_{k=i}^{N-1} x_k} \leq \frac{1 + C H_{N-1}}{1 + C}. \quad (\text{A.19})$$

Proof. We first claim that the left hand side of Eq. (A.19), as a function of \vec{x} , is maximized when $i \sum_{k=i}^{N-1} x_k \leq (i+1) \sum_{k=i+1}^{N-1} x_k$ for all $i \leq N-2$. To prove this, let \vec{x} be the maximizing vector and suppose the claim does not hold. For notational convenience, define $S_i = \sum_{k=i}^{N-1} x_k$, and let j be the smallest index such that $j S_j > (j+1) S_{j+1}$. Note by subtracting $j S_{j+1}$ from both sides, it follows that j satisfies $j x_j > S_{j+1}$. Define a new vector \vec{y} that is the same as \vec{x} except for y_j, y_{j+1} which is the solution to the following system of equations:

$$j y_j = y_{j+1} + S_{j+2}$$

$$y_j + y_{j+1} = x_j + x_{j+1}.$$

Note that this has a solution where $y_j < x_j$ and $y_{j+1} > x_{j+1}$. We shall show that \vec{y} results in a higher ratio than \vec{x} , contradicting the optimality of \vec{x} . Since $y_j + y_{j+1} = x_j + x_{j+1}$ and $y_{j+1} > x_{j+1}$, then $j y_j + (j+1) y_{j+1} > j x_j + (j+1) x_{j+1}$ which implies that the numerator of Eq. (A.19) is strictly increased under \vec{y} . Next we argue that the denominator does not change. To see this, first observe that $i \sum_{k=i}^{N-1} y_k = i \sum_{k=i}^{N-1} x_k$ for all $i \neq j+1$. Since $y_j + y_{j+1} = x_j + x_{j+1}$ and $j y_j = y_{j+1} + S_{j+2}$, then $j \sum_{k=j}^{N-1} x_k = j(y_j + y_{j+1}) + j S_{j+2} = (j+1) y_{j+1} + (j+1) S_{j+2}$ and thus the denominator is unchanged under \vec{y} . Thus \vec{y} has an increased the value of Eq. (A.19), contradicting the maximality of \vec{x} .

Thus we may assume w.l.o.g $k S_k \leq (k+1) S_{k+1}$ for any set x_1, \dots, x_{N-1} that maximizes Eq. (A.19), which in turn implies $k x_k \leq S_{k+1} \leq S_k$ for all $k \leq N$. Thus

$$\frac{1 + C \sum_{i=1}^{N-1} i x_i}{1 + C \max_i i \sum_{k=i}^{N-1} x_k} \leq \frac{1 + C \sum_{i=1}^{N-1} S_i}{1 + C \max_i i S_i} \leq \frac{1 + C S_1 H_{N-1}}{1 + C S_1} \leq \frac{1 + C H_{N-1}}{1 + C}. \quad (\text{A.20})$$

The first inequality follows from $kx_k \leq S_{k+1} \leq S_k$, the second inequality follows from Lemma A.5.1, and the third inequality from the fact that $S_1 = \sum_{k=1}^{N-1} x_i \leq 1$. \square

Lemma A.5.3. $OPT \leq 2 - \frac{1}{N}$.

Proof. First note for any optimal solution $v^* = (x, y, a, b, c, d)$ we may assume w.l.o.g. that $d = 0$ (if $d > 0$ consider v' where $(a', b', c') = (a, b, c)/(1 - d)$ and $d' = 0$). We may also scale x to 1 w.l.o.g. Then OPT is the solution of

$$\begin{aligned} & \max \frac{a + y(1 - a)}{z} \\ & \text{s.t } z \geq y \\ & z \geq a + b \\ & 0 \leq \frac{1}{N} \leq y \leq 1 \\ & a + b + c = 1 \\ & y, a, b, c \geq 0 \end{aligned}$$

At optimality either $z = y \geq a + b$ or $z = a + b \geq y$. Suppose $z = y$, then the objective becomes $1 + \frac{a - ay}{y}$, which is maximized when a is maximal. Thus the constraint $a + b \leq y$ forces $b = 0$ and $a = y$. Subbing in, then $OPT = \max_{y \geq \frac{1}{N}} 1 + (1 - y)\frac{y}{y} = 2 - \frac{1}{N}$. Similarly, suppose $z = a + b \geq y$ which implies $c \leq 1 - y$, then the objective becomes $\frac{a + y(b + c)}{1 - c}$ which is maximized when $c = 1 - y$, $a = y$, and $b = 0$. Thus $OPT = \max_{y \geq \frac{1}{N}} \frac{y + y(1 - y)}{y} = \max_{y \geq \frac{1}{N}} 2 - y = 2 - \frac{1}{N}$. \square

Lemma A.5.4. Any customer $\vec{v} \in E_2$ buys the opaque product in the 1OPQ strategy (p, p^N) .

Proof. Suppose a customer draws valuation $(v_1, v_2, \dots, v_N) \in E_2$, then $\max_i \{v_i\} =$

$p + k$, for some $k \in [0, \frac{N}{N-1}(p - p^N) - p]$. Then

$$\begin{aligned} \max_i \{v_i\} - p &= k = \frac{k}{N} + \frac{(N-1)k}{N} < \frac{k}{N} + p - p^N - p \frac{N-1}{N} \\ &= \frac{k+p}{N} - p^N = \frac{\max_i v_i}{N} - p^N \leq \frac{\sum_i v_i}{N} - p^N \end{aligned}$$

where the first inequality follows from the definition of k , and the second inequality follows from the fact that $\sum_i v_i \geq \max_i v_i$. Thus the utility from any item at price p , $\max_i \{v_i\} - p$, is less than $\frac{\sum_i v_i}{N} - p^N$, the utility from the opaque product. We also note that the utility from the opaque product is nonnegative, since $\frac{\sum_i v_i}{N} - p^N \geq \frac{k+p}{N} - p^N \geq 0$, where the last inequality follows from the case assumption $p^N < \frac{1}{N}p$. \square

Lemma A.5.5. *Let F be a distribution over m customer types. Let \vec{p} be a revenue optimal pricing and suppose that $p_1 \geq p_2 \geq \dots \geq p_N$. Then for all $i \in \mathcal{N}$, either there exists a type j such that $p_i = v_{j,i} - \max_{k>i} \{(v_{j,k} - p_k)^+\}$, or no customer type buys item i .*

Proof. Let \vec{p} be the optimal prices with $p_1 \geq \dots \geq p_N$, and let $\tilde{v}_{j,i} = v_{j,i} - \max_{k>i} (v_{j,k} - p_k)^+$. Suppose for a contradiction that there exists an item i (choose largest index if there are multiple options) that is purchased by a customer type, but $p_i \notin \{\tilde{v}_{1,i}, \dots, \tilde{v}_{m,i}\}$. Call the customer type that purchases item i as type j , and if there are multiple options select the type with smallest $\tilde{v}_{j,i}$. Note that $p_i < \tilde{v}_{j,i}$, since the reverse inequality implies type j would prefer to buy a different item (or no item) based on the definition of $\tilde{v}_{j,i}$.

We now consider an alternate pricing scheme \vec{p}' where all prices are the same except $p'_i = \tilde{v}_{j,i}$, which is a price increase. Under \vec{p}' , clearly all customers who purchased an item other than i will still purchase that item due to the price increase of i . The type j customer will buy an item with index at most i , since his favorite among the items with index greater than or equal to i is i under the new pricing (recall ties go the higher priced item), i.e., $v_{j,i} - p'_i = \max_{k>i} (v_{j,k} - p_k)^+$. Now

consider a type $l \neq j$ that also purchased i under the pricing \vec{p} . Then $v_{l,i} - p'_i = \tilde{v}_{l,i} + \max_{k>i} (v_{l,k} - p_k)^+ - p'_i \geq \tilde{v}_{j,i} + \max_{k>i} (v_{l,k} - p_k)^+ - p'_i = \max_{k>i} (v_{l,k} - p_k)^+$, where the inequality follows since j was chosen to have the smallest $\tilde{v}_{\cdot,i}$. Thus l , like j , also prefers an item with index i or lower. Therefore, \vec{p}' is a pricing with strictly better revenue, resulting in a contradiction of the optimality of \vec{p} . \square

Appendix to Chapter 2

B.1 Bounding the Value of Personalized Pricing Using the General Moments

Bounding the Value of Personalized Pricing Using Coefficient of Variation

In this subsection, we apply Theorem 2.5.2 to bound the value of personalization in terms of the *coefficient of variation* of the valuation distribution, $\frac{\sigma}{\mu} = \frac{\sqrt{E[(V-\mu)^2]}}{\mu}$. In particular, we choose $f(t) = \frac{(t-\mu)^2}{\mu^2} - \frac{\sigma^2}{\mu^2}$. Equation (2.13) then becomes

$$\max_{v \in [p_{k-1}, p_k]} \lambda_1 v + \lambda_2 \left(M^2(v-1)^2 - \frac{\sigma^2}{\mu^2} \right).$$

Since $f(\cdot)$ is convex, the possible maximizing solutions are the boundary points $\{p_{k-1}, p_k\}$ or the stationary point $1 - \frac{\lambda_1}{2M^2\lambda_2}$ if $1 - \frac{\lambda_1}{2M^2\lambda_2} \in [p_{k-1}, p_k]$ and $\lambda_2 < 0$.

We then apply the constraint generation procedure of the previous section. In Figure B.1, we plot the bound from Theorem 2.5.2 as a function of the coefficient of variation. This curve is qualitatively similar to that of Fig. 2.4, which plots the value of personalized pricing as a function of the coefficient of deviation. In both cases, the value is maximized at intermediate levels of heterogeneity. In Fig. B.1, we also show illustrate the dependence of our procedure on the choice of N , plotting the relative gap between Eq. (2.12) and the discretized version of Eq. (2.11). We observe that the

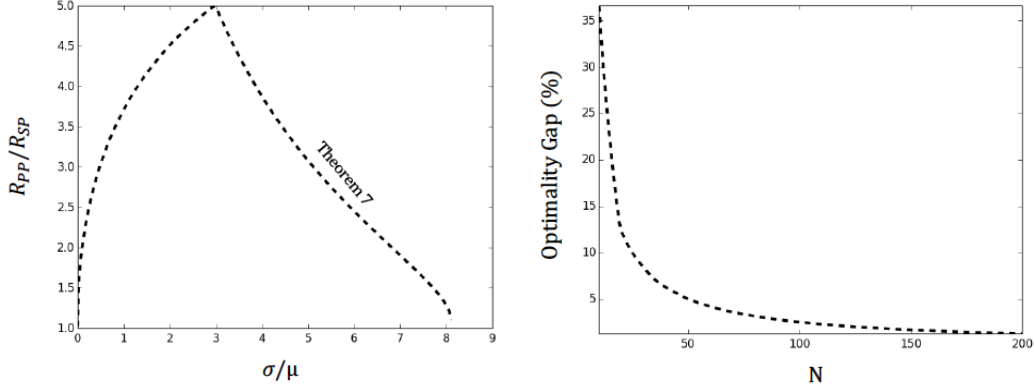


Figure B.1: The left panel shows the bound in Theorem 2.5.2 versus the variance, using with $N = 200$ discretization points. The right panel plots the percent error in Theorem 2.5.2 when $M = .9$, $S = 10$ and $\frac{\sigma}{\mu} = 4$, as a function of N , the number of discretization points.

optimality gap rapidly shrinks for relatively small values of N , suggesting that the number of samples needed to generate strong bounds on the value of personalization is not prohibitive.

Bounding the Value of Personalized Pricing Using Geometric Mean

In this subsection we apply Theorem 2.5.2 to bound the value of personalization in terms of the *geometric mean* of the valuation distribution, $G[V] := \exp(E[\log(V)])$, which was studied in [99]. As noted in [99], for any valuation V , $G[V] \leq E[V]$ and the inequality holds with equality if and only if V is a point mass. Similar to the coefficient of deviation, the statistic $\frac{G(V)}{E[V]}$ is bounded between 0 and 1 and intuitively provides a measure on the heterogeneity of F .

For simplicity, we focus on the case $c = 0$ and $\mu = 1$. (The general case can be treated by Lemma 2.3.1.) Let $f(t) = \log(t) - \log(B)$ so that $E[f(V)] = 0 \iff$

$G[V] = B$. Equation (2.13) then becomes

$$\max_{v \in [p_{k-1}, p_k]} \lambda_1 v + \lambda_2 (\log(v) - \log(B))$$

Since $f(\cdot)$ is convex, the possible maximizing solutions are the boundary points $\{p_{k-1}, p_k\}$ or the stationary point $\frac{-\lambda_2}{\lambda_1}$ if $\frac{-\lambda_2}{\lambda_1} \in [p_{k-1}, p_k]$. We then apply the constraint generation procedure of Section 2.5. In Fig. B.2(a) we plot the bound from Theorem 2.5.2 as a function of the scaled geometric mean along side Theorem 2.1 in [99]. In Fig. B.2(b) we plot the relative error in terms of N , the number of discretization points.

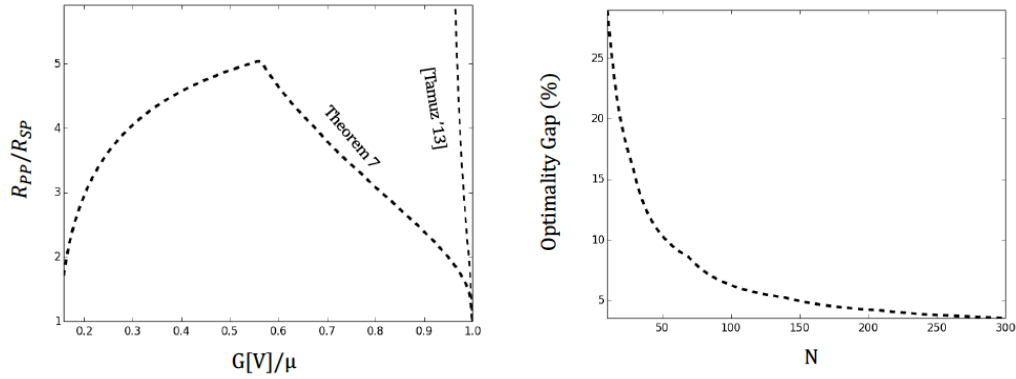


Figure B.2: The left panel shows Theorem 2.5.2 with $N = 250$ points versus the scaled geometric mean. the right panel plots the percent error in Theorem 2.5.2 when $M = .9$, $S = 10$ and $\frac{G[V]}{E[V]} = .4$, as a function of N , the number of discretization points.

Notice that for most values of $\frac{G[V]}{E[V]}$, the bound from [99] is very weak and is only useful in the limit as this quantity tends to 1. By contrast, Theorem 2.5.2 provides an essentially tight bound for all values of the parameter.

B.2 Omitted Proofs from Chapter 2

Proof of Theorem 2.3.1

The proof of Theorem 2.3.1 treats each regime of D separately. Within each regime, we utilize the same basic technique as in Lemma 2.3.2. To that end, we first establish two integral representations of D in terms of $\bar{F}(x)$.

Lemma B.2.1 (Integral Representations of D). *For any F with scale S and margin M , the coefficient of deviation D satisfies*

$$D = \int_M^{S+M-1} \bar{F}(\mu x + c) dx = \int_0^M 1 - \bar{F}(\mu x + c) dx. \quad (\text{B.1})$$

We now prove Theorem 2.3.1.

Proof. For simplicity, we first consider the special case when $c = 0$ and $\mu = 1$. In this setting $\mathcal{R}_{PP} = \mu = 1$ and $M = 1$. We follow the general technique of Lemma 2.3.2. Starting with the second identity of Lemma B.2.1,

$$D = \int_0^1 1 - \bar{F}(x) dx \geq \int_0^{\frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}} 0 dx + \int_{\frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}}^1 1 - \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{x} dx, \quad (\text{B.2})$$

where we have pointwise upper bounded $\bar{F}(x)$ by 1 for $x \in [0, \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}]$ and used the Pricing Inequality for $x \in [\frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}, 1]$. Evaluating the integrals yields,

$$D \geq \left(1 - \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}\right) + \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \log \left(\frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}\right). \quad (\text{B.3})$$

We next use properties of $W_{-1}(\cdot)$ to rewrite the inequality. For brevity, let $\alpha = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}$.

Then,

$$\begin{aligned} D \geq 1 - \alpha + \alpha \log(\alpha) &\iff D - 1 \geq \alpha(\log(\alpha) - 1) \\ &\iff \frac{D - 1}{e} \geq e^{\log(\alpha) - 1}(\log(\alpha) - 1) \quad (\text{using } \alpha = e \cdot e^{\log(\alpha) - 1}). \end{aligned}$$

Since $D \in [0, 1]$, the right hand side is between $-1/e$ and 0. Applying $W_{-1}(\cdot)$ to both sides (and recalling this function is non-increasing) yields

$$W_{-1}\left(\frac{D - 1}{e}\right) \leq \log(\alpha) - 1 \iff e \cdot e^{W_{-1}\left(\frac{D - 1}{e}\right)} \leq \alpha \quad (\text{B.4})$$

$$\iff \frac{W_{-1}\left(\frac{D - 1}{e}\right)}{D - 1} \geq \frac{1}{\alpha} \quad (\text{B.5})$$

$$\iff \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq \frac{W_{-1}\left(\frac{D - 1}{e}\right)}{D - 1}, \quad (\text{B.6})$$

where the penultimate implication follows from the definition of $W_{-1}(\cdot)$, and the last line follows from the definition of α . We stress Eq. (B.6) holds for all D and coincides with the Low Heterogeneity bound when $c = 0$, $\mu = 1$.

Similarly, we can bound the cCDF in the first identity in Lemma B.2.1 to yield an alternate bound. Specifically,

$$D = \int_1^S \bar{F}(x) dx \leq \int_1^S \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{dx}{x} = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \log(S).$$

Rearranging yields,

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq \frac{\log(S)}{D}. \quad (\text{B.7})$$

Again, we stress Eq. (B.7) holds for all D and coincides with the Medium Heterogeneity bound.

The High Heterogeneity bound can be derived similarly, using a different bounding of the cCDF which is tighter when D is large. We defer the details to the next subsection and only state the result in Lemma B.2.2 below.

Lemma B.2.2 (High Heterogeneity Bound when $c = 0$ and $\mu = 1$). *If $D > \delta_M$, then,*

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq -W_{-1} \left(\frac{-1}{eS(1-D)} \right). \quad (\text{B.8})$$

To summarize, when $c = 0$, Eqs. (B.6) and (B.7) hold for all $0 \leq D \leq \delta_H$ and Eq. (B.8) holds for all $\delta_M \leq D \leq \delta_H$. These results are sufficient to prove that the bounds from the theorem are valid. For completeness, however, the next lemma further proves that in each regime, the bound for that regime is the strongest of the applicable bounds.

Lemma B.2.3 (Strongest Bound by Regime).

a) *The function*

$$D \mapsto \frac{-W_{-1} \left(-\frac{1-D}{e} \right)}{1-D} - \frac{\log(S)}{D},$$

is negative for $D \in (0, \delta_L)$, is positive for $D \in (\delta_L, \delta_H]$, and has a unique root at $D = \delta_L$.

b) *The function*

$$D \mapsto \frac{\log(S)}{D} + W_{-1} \left(\frac{-1}{eS(1-D)} \right)$$

has a unique root at $D = \delta_M$ and is non-negative for all $D \in [0, \delta_H]$.

A consequence of Lemma B.2.3 is

- When $D \in [0, \delta_L]$, Eq. (B.6) dominates Eq. (B.7).
- When $D \in (\delta_L, \delta_M]$, Eq. (B.7) dominates Eq. (B.6).
- When $D \in (\delta_M, \delta_H]$, Eq. (B.8) dominates Eqs. (B.6) and (B.7).

This concludes the proof that the bounds are valid when $c = 0$ and $\mu = 1$.

For a general $c > 0$ and $\mu > 0$, we transform the problem to one in which $c = 0$ and $\mu = 1$ using Lemma 2.3.1 and apply the results from Eqs. (B.6) to (B.8) using the new S_c , M_c and D_c . Note, the coefficient of deviation D_c of F_c (as defined in

Lemma 2.3.1) is related to D by $D_c = D/M$. Simplifying proves that the bounds are valid for general c and μ .

It only remains to establish that the bounds are tight. We use the same technique as in Lemma 2.3.2. Namely, in each regime, given S , M , D , and μ , we construct a cCDF that makes all pointwise bounds on the cCDF simultaneously. A difference from Lemma 2.3.2 is that the integral representations of D in the proof of Theorem 2.3.1 do not determine \bar{F} over its whole domain $[0, S\mu]$; they only span $[0, \mu]$, or $[\mu, S]$ depending on the regime. This introduces some freedom in constructing the cCDF on the remaining segment and causes the tight distributions to be non-unique. Nonetheless, since these constructions follow the proof of Lemma 2.3.2 closely, we defer the details to Lemma B.2.4 in the next subsection for brevity. \square

Omitted Proofs from Theorem 2.3.1.

We now provide proofs for the lemmas necessary to complete the proof of Theorem 2.3.1.

Proof of Lemma B.2.1. Let $V \sim F$ and note,

$$0 = E[V - \mu] = E[(V - \mu)^+] - E[(\mu - V)^+] \implies E[(V - \mu)^+] = E[(\mu - V)^+].$$

Moreover, $E[|V - \mu|] = E[(V - \mu)^+] + E[(\mu - V)^+]$, hence, combining with the above yields $E[|V - \mu|] = 2E[(V - \mu)^+] = 2E[(\mu - V)^+]$. We use these two identities to re-express D . From the first equality and the tail integral formula for expectation,

$$\begin{aligned} D &= \frac{1}{\mu} E[(V - \mu)^+] = \frac{1}{\mu} \int_0^\infty \Pr((V - \mu)^+ \geq t) dt \\ &= \frac{1}{\mu} \int_0^{\mu(S-1)} \Pr(V \geq \mu + t) dt = \int_M^{S+M-1} \bar{F}(\mu x + c) dx, \end{aligned}$$

where the last line follows from the change of variables $\mu + t \rightarrow \mu x + c$. Similarly,

using second equality and the tail integral formula for expectation,

$$\begin{aligned} D &= \frac{1}{\mu} E[(\mu - V)^+] = \frac{1}{\mu} \int_0^\infty \Pr((\mu - V)^+ > t) dt \\ &= \frac{1}{\mu} \int_0^{\mu-c} \Pr(V \leq \mu - t) dt = \int_0^M F(\mu x + c) dx \end{aligned}$$

where the last line follows from the change of variables $\mu - t \rightarrow \mu x + c$. \square

Proof of Lemma B.2.2. We follow the same strategy as previous two regimes bounds. Note that when the coefficient of deviation is high, the probability that V is “close” to 1 is low, since $\mu = 1$. Formally, we claim that

$$\Pr(V \geq t) \leq 1 - D \quad \forall t \in (1, S) \quad (\text{B.9})$$

To prove the claim, note that $D = E[(1 - V)^+] \leq \Pr(V \leq 1)$, where the equality is Lemma B.2.1 and the inequality uses $(1 - V)^+ \leq 1$. Rearranging proves $\Pr(V \geq 1) \leq 1 - D$, which in turn implies Eq. (B.9).

We use this inequality when pointwise bounding our integral representation. Specifically, for any $1 \leq t_0 \leq S$, we have

$$\begin{aligned} D &= \int_1^S \Pr(V > t) dt && (\text{Lemma B.2.1}) \\ &= \int_1^{t_0} \Pr(V > t) dt + \int_{t_0}^S \Pr(V > t) dt \\ &\leq \int_1^{t_0} (1 - D) dt + \int_{t_0}^S \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{dt}{t} && (\text{Eq. (B.9) and Pricing Inequality}) \\ &= (t_0 - 1)(1 - D) + \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \log \left(\frac{S}{t_0} \right) && (\text{B.10}) \end{aligned}$$

Minimizing over t_0 yields $t_0 = \max \left\{ 1, \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)} \right\}$. We next argue that $D \geq \delta_M$ implies $1 \leq \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)}$, so that the unique minimizer is $t_0 = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)}$.

Recall by Eq. (B.7) $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq \frac{\log(S)}{D}$ for all values of D and, in particular, we have that for $D \in [\delta_M, \delta_H]$,

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq \frac{\log(S)}{D} \leq \frac{\log(S)}{\delta_M} = 1 + \log(S).$$

Further $D \geq \delta_M = \frac{\log(S)}{1+\log(S)}$ implies that $1 + \log(S) \leq \frac{1}{1-D}$. Combining shows

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq \frac{1}{1-D} \iff 1 \leq \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)},$$

which confirms that $t_0 = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)}$ is the unique minimizer.

Plugging in this value $t_0 = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \frac{1}{(1-D)}$ into Eq. (B.10) yields:

$$1 \leq \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} + \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}} \log \left(\frac{S(1-D)}{\frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}} \right)$$

We next use properties of the Lambert- W function to simplify this equation. For notational convenience define $\alpha = \frac{\mathcal{R}_{SP}}{\mathcal{R}_{PP}}$. Then,

$$\begin{aligned} 1 \leq \alpha + \alpha \log \left(\frac{S(1-D)}{\alpha} \right) &\iff 1 \leq \alpha(1 + \log(S(1-D)) - \log(\alpha)) \quad (\text{B.11}) \\ &\iff -1 \geq \alpha(\log(\alpha) - \log(eS(1-D))) \end{aligned}$$

Note $\alpha = e^{\log(\alpha)} = e^{\log(\alpha) - \log(eS(1-D))} \cdot e \cdot S(1-D)$. Substituting above proves

$$\frac{-1}{eS(1-D)} \geq e^{\log(\alpha) - \log(eS(1-D))} (\log(\alpha) - \log(eS(1-D))).$$

The left hand side is between $-1/e$ and 0 by inspection. The function $W_{-1}(\cdot)$ is

non-increasing on this range, so that applying $W_{-1}(\cdot)$ to both sides yields

$$\begin{aligned}
W_{-1}\left(\frac{-1}{eS(1-D)}\right) &\leq \log(\alpha) - \log(eS(1-D)) & (B.12) \\
\iff \alpha &\geq eS(1-D) \cdot e^{W_{-1}\left(\frac{-1}{eS(1-D)}\right)} \\
\iff \frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} &\leq -\frac{1}{eS(1-D)} e^{-W_{-1}\left(\frac{-1}{eS(1-D)}\right)}.
\end{aligned}$$

Finally, from the definition of W_{-1} ,

$$\frac{-1}{eS(1-D)} = W_{-1}\left(\frac{-1}{eS(1-D)}\right) e^{W_{-1}\left(\frac{-1}{eS(1-D)}\right)},$$

which we use to simplify the last inequality to obtain $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \leq -W_{-1}\left(\frac{-1}{eS(1-D)}\right)$. \square

Proof of Lemma B.2.3. First consider part *a*). Recalling that $-W_{-1}(-1/e) = 1$, we confirm directly that the given function is negative as $D \downarrow 0$ since it is continuous. Notice further that $-W_{-1}(\cdot)$ is an increasing function (cf. Fig. 2.2), whereby $\frac{-W_{-1}\left(-\frac{1-D}{e}\right)}{1-D}$ is an increasing function, while $\log(S)/D$ is a decreasing function. It follows that the given function has a unique root, and it suffices to show this root is δ_L to complete the proof. To this end, write,

$$\begin{aligned}
-\frac{W_{-1}\left(-\frac{1-D}{e}\right)}{1-D} = \frac{\log(S)}{D} &\iff W_{-1}\left(-\frac{1-D}{e}\right) = \log\left(S^{\frac{D-1}{D}}\right) \\
&\iff -\frac{1-D}{e} = \log\left(S^{\frac{D-1}{D}}\right) \cdot \exp\left(\log\left(S^{\frac{D-1}{D}}\right)\right) \\
&\iff -\frac{1}{eS} = S^{\frac{-1}{D}} \cdot \frac{-\log(S)}{D} \\
&\iff -\frac{1}{eS} = \exp\left(-\frac{\log(S)}{D}\right) \cdot \frac{-\log(S)}{D} \\
&\iff W_{-1}\left(-\frac{1}{eS}\right) = -\frac{\log(S)}{D} \\
&\iff D = -\frac{\log(S)}{W_{-1}\left(-\frac{1}{eS}\right)} = \delta_L.
\end{aligned}$$

The second equivalence follows from the definition of the Lambert-W function. The third follows by simplifying. The fourth follows by noting $S^{\frac{-1}{D}} = \exp\left(-\frac{\log(S)}{D}\right)$. The first follows from applying $W_{-1}(\cdot)$. This completes the proof of part a).

To prove part b), first observe that

$$W_{-1}\left(-\frac{1}{eS(1-D)}\right) \geq -\frac{\log(S)}{D} \iff -\frac{1}{eS(1-D)} \leq -\frac{\log(S)}{D} \exp\left(-\frac{\log(S)}{D}\right),$$

because the function $y \mapsto ye^y$ is the inverse of $W_{-1}(\cdot)$ and is non-increasing on the domain of $W_{-1}(\cdot)$, i.e., $[-1/e, 0)$. Simplifying the righthand inequality yields,

$$\frac{-1}{e} \leq \log\left(S^{\frac{D-1}{D}}\right) \cdot S^{\frac{D-1}{D}}.$$

Now make the substitution $\log\left(S^{\frac{D-1}{D}}\right) \rightarrow y$ so this last inequality is equivalent to $\frac{-1}{e} \leq ye^y$. One can confirm by differentiation that $y \mapsto ye^y$ has a unique minimizer at $y = -1$, and, thus, this last inequality holds for all y . This proves the function defined in part b) is nonnegative everywhere. Moreover, it has a root at $y = 1$ which corresponds to $\log\left(S^{\frac{D-1}{D}}\right) = -1$. Simplifying shows this condition is equivalent to $D = \log(S)/(1 + \log(S)) = \delta_M$, as was to be proven. \square

We next explicitly describe the distributions which make Theorem 2.3.1 tight. By Lemma 2.3.1, it suffices to consider the case where $c = 0$ and $\mu = 1$. The general case can be handled by scaling and shifting the below tight distributions:

Lemma B.2.4 (Tight distributions).

a) Suppose $D \in [0, \delta_L]$, and let $\alpha_L = \left(\frac{W_{-1}\left(\frac{D-1}{e}\right)}{D-1}\right)^{-1}$. Then, there is a random

variable V with $cCDF$

$$\bar{F}_L(x) = \begin{cases} 1 & \text{if } 0 \leq x < \alpha_L \\ \frac{\alpha_L}{x} & \text{if } \alpha_L \leq x \leq 1 \\ \frac{D}{\log(S)x} & \text{if } 1 < x \leq S \\ 0 & \text{otherwise,} \end{cases} \quad (\text{Tight cCDF, Low Heterogeneity})$$

and this random variable has scale S , coefficient of deviation D , and mean 1 and satisfies Eq. (B.6) with equality.

b) Suppose $D \in [\delta_L, \delta_M]$, and let $\alpha_M = \frac{D}{\log(S)}$. Then, there is a random variable V with $cCDF$

$$\bar{F}_M(x) = \begin{cases} 1 & \text{if } x = 0, \\ \frac{\alpha_M}{e} S^{\frac{1}{D}-1} & \text{if } x \in (0, eS^{1-\frac{1}{D}}) \\ \frac{\alpha_M}{x} & \text{if } x \in [eS^{1-\frac{1}{D}}, S] \\ 0 & \text{otherwise,} \end{cases} \quad (\text{Tight cCDF, Medium Heterogeneity})$$

and this random variable has scale S , coefficient of deviation D , and mean 1 and satisfies Eq. (B.7) with equality.

c) Suppose $D \in [\delta_M, \delta_H]$, and let $\alpha_H := \left(-W_{-1}\left(\frac{-1}{eS(1-D)}\right)\right)^{-1}$. Then, there is a random variable V with $cCDF$

$$\bar{F}_H(x) = \begin{cases} 1 & \text{if } x = 0, \\ 1 - D & \text{if } x \in (0, \frac{\alpha_H}{1-D}] \\ \frac{\alpha_H}{x} & \text{if } x \in (\frac{\alpha_H}{1-D}, S) \\ 0 & \text{otherwise,} \end{cases} \quad (\text{Tight cCDF, High Heterogeneity})$$

and this random variable has scale S , coefficient of deviation D , and mean 1

and satisfies Eq. (B.8) with equality.

Proof of Lemma B.2.4. Intuitively, \bar{F}_L , \bar{F}_M , and \bar{F}_H each make all the pointwise bounds on the cCDF the integral representation of D used in the proofs of Eqs. (B.6) to (B.8) tight, simultaneously. Thus, they will make the overall bound tight.

To prove the lemma formally, we will prove that \bar{F}_L , \bar{F}_M and \bar{F}_H are valid cCDFs, each with mean 1, scale S , and coefficient of deviation D , and that $\mathcal{R}_{SP}(F_L, 0) = \alpha_L$, $\mathcal{R}_{SP}(F_M, 0) = \alpha_M$ and $\mathcal{R}_{PP}(F_H, 0) = \alpha_H$, respectively. The lemma then follows directly from the definition of α_L , α_M and α_H since $\mathcal{R}_{PP}(F_L, 0) = \mathcal{R}_{PP}(F_M, 0) = \mathcal{R}_{PP}(F_H, 0) = \mu = 1$.

a) (Low Heterogeneity) Note that replacing α by α_L and the inequality by equality in Eq. (B.4) and then following the implications backwards proves that α_L satisfies

$$D = 1 - \alpha_L + \alpha_L \log(\alpha_L).$$

We next prove \bar{F}_L is a valid cCDF. By inspection, we need only prove \bar{F}_L is non-increasing, i.e., that $\alpha_L \geq D/\log(S) \iff 1/\alpha_L \leq \log(S)/D$. This inequality follows directly from Lemma B.2.3 since $D \in [0, \delta_L]$, and the lefthand side is low-heterogeneity bound while the right side is the medium heterogeneity bound. This proves \bar{F}_L is valid.

Next, write

$$\int_0^\infty \bar{F}_L(t) dt = \int_0^1 \bar{F}_L(x) dx + \int_1^S \bar{F}_L(x) dx = \alpha_L - \alpha_L \log(\alpha_L) + D = 1,$$

where the last equality uses the identity proven above for α_L . Thus, \bar{F}_L has mean 1. By Lemma B.2.1, its coefficient of deviation is

$$\int_0^1 1 - \bar{F}_L(x) dx = \int_0^{\alpha_L} 0 dx + \int_{\alpha_L}^1 1 - \frac{\alpha_L}{x} dx = 1 - \alpha_L + \alpha_L \log(\alpha_L) = D, \quad (\text{B.13})$$

again using the identify for α_L . By inspection, it has scale S .

Finally, any price $x \in [\alpha_L, 1]$ earns profit α_L , while any price $x \in [0, \alpha_L)$ earns profit strictly less than α_L . Any price $x \in (1, S]$ earns profit $D/\log(S)$ which is at most α_L as we noted when proving that \bar{F}_L is valid. Thus, $\mathcal{R}_{SP}(F_L, 0) = \alpha_L$, which proves that a random variable V with cCDF \bar{F}_L will satisfy Eq. (B.6) with equality.

b) (Medium Heterogeneity) To prove that \bar{F}_M is a valid cCDF, it suffices to show that $eS^{1-\frac{1}{D}} \leq S$, which is equivalent to $1 \geq \frac{D}{\log(S)}$. Rewrite this last inequality as $\frac{1}{\alpha_M} \geq 1$, and recall from Step 1 of the proof of Theorem 2.3.1 that $\frac{1}{\alpha_M}$ is an upper bound on the value of personalization and, thus, must be at least 1.

Next, write

$$\int_0^\infty \bar{F}_M(x)dx = \int_0^{eS^{1-\frac{1}{D}}} \bar{F}_M(x)dx + \int_{eS^{1-\frac{1}{D}}}^S \bar{F}_M(x)dx = \alpha_M + \alpha_M \log\left(\frac{S}{eS^{1-\frac{1}{D}}}\right) = 1,$$

where the last equality uses the definition of α_M . It follows that \bar{F}_M has mean 1, and, by inspection, scale S . Write,

$$\int_1^S \bar{F}_M(x)dx = \alpha_M \log S = D,$$

to conclude from Lemma B.2.1 that \bar{F}_M has coefficient of deviation D . Finally, observe that any price $x \in [eS^{1-\frac{1}{D}}, S]$ earns profit α_M , while any other price earns strictly less profit. Thus, $\mathcal{R}_{SP}(F_M, 0) = \alpha_M$, completing this part of the lemma.

c) (High Heterogeneity) To prove \bar{F}_H is a valid cCDF, it suffices to show that $\alpha_H/(1-D) \leq S$. Note that by Lemma B.2.2, $1/\alpha_H$ is an upper-bound on the value of personalization, whereby α_H is necessarily at most 1. Moreover, for the Lambert- W function defining α_H to be well-defined, we must have that $\frac{1}{S(1-D)} \leq 1$ which implies $S(1-D) \geq 1$. Thus, $\alpha_H \leq 1 \leq S(1-D)$ which implies that $\alpha_H/(1-D) \leq S$ and

that \bar{F}_H is a valid cCDF.

Next write,

$$\begin{aligned} \int_0^S \bar{F}(x) dx &= \int_0^{\frac{\alpha_H}{1-D}} (1-D) dx + \int_{\frac{\alpha_H}{1-D}}^S \frac{\alpha_H}{x} dx \\ &= \alpha_H + \alpha_H \log \left(\frac{S}{\alpha_H} (1-D) \right). \end{aligned} \quad (\text{B.14})$$

We claim this last quantity equals 1. Indeed, from the definition of $W_{-1}(\cdot)$, $\alpha_H = eS(1-D) \cdot e^{W_{-1}(\frac{-1}{eS(1-D)})}$. Then, replace α by α_H and the inequality by equality in Eq. (B.12) and follow the implications backwards to Eq. (B.11), proving the claim. Thus, \bar{F}_H has mean 1, and, by inspection, has scale S .

To compute its coefficient of deviation, we first claim that $\alpha_H/(1-D) \geq 1$. Indeed, recall that

$$D \geq \delta_M = \frac{\log(S)}{1 + \log(S)} \iff \log(S) \leq \frac{D}{1-D} \iff \frac{\log(S)}{D} \leq \frac{1}{1-D}.$$

It follows that

$$\frac{\alpha_H}{1-D} \geq \alpha_H \frac{\log(S)}{D} = \frac{\alpha_H}{\alpha_M} \geq 1,$$

where the last inequality follows from Lemma B.2.3. Now compute

$$\int_0^1 1 - \bar{F}_H(x) dx = D,$$

whereby \bar{F}_H has coefficient of deviation D by Lemma B.2.1.

$$\int_0^1 1 - \bar{F}(x) dx = D.$$

It remains to check that $\mathcal{R}_{SP}(F, 0) = \alpha_H$, which we verify directly by observing that any price $x \in [\frac{\alpha_H}{1-D}, S]$ obtains profit α_H any other price obtains profit no more than α_H . □

Other Omitted Proofs from Section 2.3

Proof of Lemma 2.3.1. First note the profit from personalized pricing under valuation distribution F is $\mathcal{R}_{PP}(F, c) = E[V] - c = \mu - c$ and under F_c is $\mathcal{R}_{PP}(F_c, 0) = E[\frac{1}{\mu-c}(V - c)] - 0 = 1$. Hence, it suffices to show that $\mathcal{R}_{SP}(F, c) = (\mu - c)\mathcal{R}_{SP}(F_c, 0)$ to prove the first statement. Observe that

$$\begin{aligned}
 \mathcal{R}_{SP}(F, c) &= \max_p (p - c) \Pr(V \geq p) \\
 &= \max_p (p - c) \Pr\left(\frac{V - c}{\mu - c} \geq \frac{p - c}{\mu - c}\right) \\
 &= \max_q (\mu - c)q \Pr\left(\frac{V - c}{\mu - c} \geq q\right) \quad \left(\text{Making the substitution } \frac{p - c}{\mu - c} \rightarrow q\right) \\
 &= (\mu - c)\mathcal{R}_{SP}(F_c, 0).
 \end{aligned}$$

For the last statement of the theorem, note that $\mu_c = E[\frac{1}{\mu-c}(V - c)] = 1$, $M_c = 1 - 0/\mu_c = 1$, and

$$\begin{aligned}
 S_c &= \frac{\inf\{k \mid F_c(k) = 1\}}{\mu_c} = \frac{\frac{1}{\mu-c}(\inf\{k \mid F(k) = 1\} - c)}{1} \\
 &= \frac{\mu}{\mu - c} \left(\frac{\inf\{k \mid F(k) = 1\}}{\mu} - \frac{c}{\mu} \right) = \frac{S - 1 + M}{M}.
 \end{aligned}$$

This completes the proof. \square

Proof of Lemma 2.3.3. Consider the case when $c = 0$ and $\mu = 1$, which implies that $M = 1$. We first prove that $D \leq \delta_H$ and that there exists an F whose coefficient of deviation is exactly δ_H . To this end, consider an arbitrary random variable V , and define the new random variable \bar{V} with two-point support

$$\bar{V} = \begin{cases} E[V|V \leq 1] & \text{with probability } \Pr(V \leq 1) \\ E[V|V > 1] & \text{with probability } \Pr(V > 1). \end{cases}$$

By construction, $E[\bar{V}] = E[V] = 1$. Furthermore,

$$\begin{aligned}
E[|V - 1|] &= E[|V - 1| \mid V \leq 1] \Pr(V \leq 1) + E[|V - 1| \mid V > 1] \Pr(V > 1) \\
&= E[1 - V \mid V \leq 1] \Pr(V \leq 1) + E[V - 1 \mid V > 1] \Pr(V > 1) \\
&= \left(1 - E[V \mid V \leq 1]\right) \Pr(V \leq 1) + \left(E[V \mid V > 1] - 1\right) \Pr(V > 1) \\
&= E[|\bar{V} - 1|],
\end{aligned}$$

i.e., both V and \bar{V} have the same coefficient of deviation. Thus, to find a distribution with maximal coefficient of deviation, it suffices to consider two-point distributions.

We compute such a distribution explicitly via the following optimization problem:

$$\begin{aligned}
&\frac{1}{2} \max_{x,y,q} \quad q(1-x) + (1-q)(y-1) \\
&\text{s.t.} \quad qx + (1-q)y = 1 \\
&\quad 0 \leq x \leq 1 \leq y \leq S, \quad 0 \leq q \leq 1,
\end{aligned}$$

where the objective is the coefficient of deviation of a distribution with mass q at $x < 1$ and mass $1-q$ at $y > 1$. The constraint ensures that the mean is 1. In particular, this constraint implies $q = \frac{y-1}{y-x}$ for any feasible solution, whereby the objective simplifies to $\frac{(1-x)(2y-1)}{y-x}$. This function is decreasing in x , whereby the optimal solution is $x^* = 0$, $y^* = S$ and $q^* = \frac{S-1}{S}$ with optimal value $\frac{S-1}{S}$. Note $\frac{S-1}{S} = \delta_H$ since $M = 1$.

Next we show $0 \leq \delta_L \leq \delta_M \leq \delta_H$. Notice that $\delta_L = \frac{\log(S)}{-W_{-1}(-\frac{1}{eS})}$ is the ratio of two positive terms. Thus, it is positive. To show $\delta_L \leq \delta_M$, note that, since $S \geq 1$,

$$1 + \log(S) \geq 1 = \frac{e^{1+\log(S)}}{eS},$$

which, after rearranging, implies

$$-(1 + \log(S)) e^{-(1+\log(S))} \leq \frac{-1}{eS}.$$

Applying $W_{-1}(\cdot)$ to both sides and noting this function is decreasing shows

$$-(1 + \log(S)) \geq W_{-1}\left(\frac{-1}{eS}\right),$$

which implies

$$\delta_L = \frac{\log(S)}{-W_{-1}\left(\frac{-1}{eS}\right)} \leq \frac{\log(S)}{1 + \log(S)} = \delta_M,$$

as was to be shown.

To show $\delta_M \leq \delta_H$, observe that since $S \geq 1$, $0 \leq \log(S) \leq S - 1$, which implies that

$$\delta_M = \frac{\log(S)}{1 + \log(S)} \leq \frac{S - 1}{1 + (S - 1)} = \delta_H,$$

since $x \mapsto \frac{x}{1+x}$ is an increasing function for $x \geq 0$. This completes the proof in the case $c = 0$ and $\mu = 1$.

For general $c > 0$ and $\mu > 0$, first apply Lemma 2.3.1 to obtain an instance with zero cost and unit mean with corresponding parameters D_c, S_c , and M_c . From the previous arguments, we have that $0 \leq D_c \leq \frac{S_c - 1}{S_c}$ and $0 \leq \frac{\log(S_c)}{W_{-1}\left(\frac{-1}{eS_c}\right)} \leq \frac{\log(S_c)}{1 + \log(S_c)} \leq \frac{S_c - 1}{S_c}$. Transform back to the original parameters to prove the lemma, noting that $D_c = \frac{D}{M}$ and $S_c = \frac{S + M - 1}{M}$. \square

Proof of Lemma 2.3.4. Let us fix S and M , and define $\alpha(D) := \alpha(S, M, D)$. Fix any D_1, D_2 , with $0 \leq D_1 \leq D_2 \leq \delta_H$, and any $t \in [0, 1]$. We will show that $\alpha(tD_1 + (1 - t)D_2) \leq t\alpha(D_1) + (1 - t)\alpha(D_2)$ to prove the theorem.

By Theorem 2.3.1, there exists random variables $V_1 \sim F_1$ and $V_2 \sim F_2$ each with scale S and margin M such that the coefficient of deviation of F_1 is D_1 , the coefficient

of deviation of F_2 is D_2 , $\alpha(D_1) = \frac{\mathcal{R}_{SP}(F_1, c)}{\mathcal{R}_{PP}(F_1, c)}$ and $\alpha(D_2) = \frac{\mathcal{R}_{SP}(F_2, c)}{\mathcal{R}_{PP}(F_2, c)}$.

Since both V_1 and V_2 have the same margin and cost, they also have the same mean $\mu = \frac{c}{1-M}$. Take X to be a Bernoulli random variable with parameter t , and let $\tilde{V} \equiv XV_1 + (1 - X)V_2$ where X, V_1, V_2 are sampled independently. Note that \tilde{V} has mean μ , margin M , and scale S . Furthermore, the coefficient of deviation of \tilde{V} is

$$\begin{aligned}\tilde{D} &= \frac{1}{2\mu} \left(E \left[|XV_1 + (1 - X)V_2 - \mu| \right] \right) \\ &= \Pr(X = 1) \cdot \frac{1}{2\mu} E \left[|V_1 - \mu| \right] + \Pr(X = 0) \cdot \frac{1}{2\mu} E \left[|V_2 - \mu| \right] \\ &= tD_1 + (1 - t)D_2.\end{aligned}\tag{B.15}$$

To conclude the proof, write

$$\begin{aligned}t\alpha(D_1) + (1 - t)\alpha(D_2) &= t \frac{\mathcal{R}_{SP}(F_1, c)}{\mathcal{R}_{PP}(F_1, c)} + (1 - t) \frac{\mathcal{R}_{SP}(F_2, c)}{\mathcal{R}_{PP}(F_2, c)} \\ &= \frac{t\mathcal{R}_{SP}(F_1, c) + (1 - t)\mathcal{R}_{SP}(F_2, c)}{\mathcal{R}_{PP}(\tilde{F}, c)} \\ &\geq \frac{\mathcal{R}_{SP}(\tilde{F}, c)}{\mathcal{R}_{PP}(\tilde{F}, c)} \\ &\geq \alpha(\tilde{D}) \\ &= \alpha(tD_1 + (1 - t)D_2).\end{aligned}$$

The first equation follows from the definitions of F_1 and F_2 . The second equation follows from the fact that the personalized pricing strategy yields $\mu - c$ for F_1, F_2 , and \tilde{F} . The first inequality follows from the fact that the optimal single price for \tilde{V} yields revenue of at most $\mathcal{R}_{SP}(F_1, c)$ for the market corresponding to V_1 and at most $\mathcal{R}_{SP}(F_2, c)$ for the market corresponding to V_2 . The second inequality follows Theorem 2.3.1. The last equality follows from Eq. (B.15). \square

Proof of Corollary 2.3.1. Note that Eq. (2.1) shows that

$$W_{-1}\left(-\frac{x}{e}\right) = 1 + \sqrt{2\log(1/x)} + O(\log(1/x)) \quad \text{as } x \rightarrow 1.$$

Substituting this expression into the bounds in the low heterogeneity and high heterogeneity regimes proves the result. \square

Proof of Theorem 2.3.2. For simplicity, we consider the special case when $c = 0$, the general case is handled by transforming via Lemma 2.3.1. We will follow the geometric intuition depicted in Fig. 2.6(a). First we will construct a trapezoidal lower bound on $\mathcal{R}_{PP}(F, 0)$, then an upper bound $\mathcal{R}_{SP}(F, 0)$ by considering the area of any inscribed rectangle. Fix F and let α be the mode of F . Define $\lambda := \frac{\int_0^\alpha \bar{F}(x)dx}{\mu}$ and suppose the revenue maximizing single price p^* is less than α , then it follows that $\mathcal{R}_{SP}(F, 0) \leq \int_0^{p^*} \bar{F}(x)dx \leq \lambda\mu$ which in turn implies $\frac{\mathcal{R}_{PP}(F, 0)}{\mathcal{R}_{SP}(F, 0)} \geq \frac{1}{\lambda}$. For the remainder of the proof suppose that $p^* > \alpha$ and let $p^* = (1 + k)\alpha$ for some $k > 0$.

By unimodality, \bar{F} is convex on $[\alpha, \infty)$, thus there is a supporting line at p^* which we will denote as $l_{p^*}(x) := \bar{F}(p^*) - f(p^*)(x - p^*)$. Since $\bar{F}(x)$ is decreasing, for any $x \in [0, \alpha]$, $\bar{F}(x) \geq \bar{F}(\alpha) \geq l_{p^*}(\alpha) = \bar{F}(p^*) - f(p^*)(\alpha - p^*)$. Integrating yields a lower bound on \mathcal{R}_{PP} ,

$$\begin{aligned} \int_0^\infty \bar{F}(x)dx &\geq \int_0^{p^* + \frac{\bar{F}(p^*)}{f(p^*)}} \bar{F}(x)dx = \int_0^\alpha \bar{F}(x)dx + \int_\alpha^{p^* + \frac{\bar{F}(p^*)}{f(p^*)}} \bar{F}(x)dx \\ &\geq \alpha l_{p^*}(\alpha) + l_{p^*}(\alpha) \frac{\left(p^* + \frac{\bar{F}(p^*)}{f(p^*)}\right) - \alpha}{2} \\ &= l_{p^*}(\alpha) \left(\frac{\alpha}{2} + p^*\right), \end{aligned} \tag{B.16}$$

where the final equality follows from the first order optimality conditions for the revenue maximizing price, $\frac{d}{dp} p\bar{F}(p^*) = 0 \implies \bar{F}(p^*) = f(p^*)p^*$. We will now compare $\mathcal{R}_{SP}(F, 0)$ with Eq. (B.16).

By definition of the supporting line at p^* , $l_{p^*}(p^*) = \bar{F}(p^*)$ and the line has a unique root $l_{p^*}(p^* + \frac{\bar{F}(p^*)}{f(p^*)}) = l_{p^*}(2p^*) = 0$. Using these two points we may rewrite the supporting line as $l_{p^*}(x) = \frac{2l_{p^*}(\alpha)p^* - l_{p^*}(\alpha)x}{2p^* - \alpha}$ from which we can derive that $\bar{F}(p^*) = l_{p^*}(p^*) = \frac{l_{p^*}(\alpha)p^*}{2p^* - \alpha}$. Combining this expression with Eq. (B.16) we obtain

$$\frac{\mathcal{R}_{PP}(F, 0)}{\mathcal{R}_{SP}(F, 0)} \geq \frac{l_{p^*}(\alpha) \left(\frac{\alpha}{2} + p^*\right)}{p^* \frac{l_{p^*}(\alpha)p^*}{2p^* - \alpha}} = 2 - \frac{1}{2} \left(\frac{\alpha}{p^*}\right)^2. \quad (\text{B.17})$$

To complete the proof will proceed in two cases depending on the size of λ .

(Case 1: $\lambda \geq \frac{2}{3}$) By rearranging the equality Eq. (B.17) we can rewrite \mathcal{R}_{SP} as,

$$\begin{aligned} \mathcal{R}_{SP}(F, 0) &= \frac{1}{2 - \frac{1}{2} \left(\frac{\alpha}{p^*}\right)^2} \cdot l_{p^*}(\alpha) \left(\frac{\alpha}{2} + p^*\right) && (\text{Eq. (B.17)}) \\ &\leq \frac{1}{2 - \frac{1}{2} \left(\frac{\alpha}{p^*}\right)^2} \cdot l_{p^*}(\alpha) \cdot 3 \left(\frac{2p^* - \alpha}{2}\right) && (\alpha \leq p^*) \\ &\leq 2 \left(l_{p^*}(\alpha) \frac{2p^* - \alpha}{2}\right) && (\alpha \leq p^*) \\ &\leq 2(1 - \lambda) \leq \lambda. \end{aligned}$$

The third inequality follows from noting $1 - \lambda = \int_{\alpha}^{2p^*} \bar{F} dx \geq \int_{\alpha}^{2p^*} l_{p^*}(x) = l_{p^*}(\alpha) \frac{2p^* - \alpha}{2}$.

The final inequality follows from $\lambda \geq 2/3$.

(Case 2: $\lambda \leq \frac{2}{3}$) Following the geometry in Fig. 2.6(a), first write $\mathcal{R}_{SP}(F, 0)$ as the sum of the area before the mode and after the mode i.e., $\mathcal{R}_{SP}(F, 0) = \alpha \bar{F}(p^*) + (p^* - \alpha) \bar{F}(p^*)$. The first term in the sum is bounded by $\alpha \bar{F}(p^*) = \alpha \bar{F}(\alpha) \frac{\alpha \bar{F}(p^*)}{\alpha \bar{F}(\alpha)} \leq \lambda \frac{\alpha \bar{F}(p^*)}{\alpha \bar{F}(\alpha)} = \lambda \frac{p^*}{2p^* - \alpha}$. The second term is bounded by $(p^* - \alpha) \bar{F}(p^*) = l_{p^*}(\alpha)(p^* -$

$\frac{\alpha}{2}) \frac{(p^* - \alpha) \bar{F}(p^*)}{l_{p^*}(\alpha)(p^* - \frac{\alpha}{2})} \leq (1 - \lambda) \frac{(p^* - \alpha) \bar{F}(p^*)}{l_{p^*}(\alpha)(p^* - \frac{\alpha}{2})} = (1 - \lambda) \frac{2p^*(p^* - \alpha)}{(2p^* - \alpha)^2}$. Then,

$$\begin{aligned} \mathcal{R}_{SP}(F, 0) &\leq \lambda \frac{p^*}{2p^* - \alpha} + (1 - \lambda) \frac{2p^*(p^* - \alpha)}{(2p^* - \alpha)^2} \\ &= \frac{1 + k}{1 + 2k} \left(\lambda + \frac{2k}{1 + 2k} (1 - \lambda) \right) \quad (p^* = (1 + k)\alpha \text{ for some } k > 0) \\ &\leq \frac{(\lambda - 2)^2}{8(1 - \lambda)} \quad \left(\text{Maximized when } k = \frac{1}{\lambda} - \frac{3}{2} \right) \end{aligned}$$

Thus when $\lambda \leq 2/3$, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{SP}} \geq \min\{\frac{1}{\lambda}, \frac{8(1-\lambda)}{(\lambda-2)^2}\} = \frac{8(1-\lambda)}{(\lambda-2)^2}$. To complete the proof we will use Lemma B.2.1 to relate λ and D by

$$\lambda = \int_0^\alpha \bar{F}(x) dx = \int_0^\mu \bar{F}(x) dx + \int_\mu^\alpha \bar{F}(x) dx = 1 - D + \int_\mu^\alpha \bar{F}(x) dx \leq 1 - D.$$

Substituting in both cases gives the result. \square

Omitted Proofs from Section 2.4

Proof of Theorem 2.4.1. We let $F_{V|\mathbf{X}}$ denote the conditional distribution of $V \mid \mathbf{X}$, with corresponding mean $\mu(\mathbf{X})$, scale $S(\mathbf{X})$, margin $M(\mathbf{X})$, and coefficient of deviation $D(\mathbf{X})$. By assumption, $M(\mathbf{X}) = 1 - \frac{c}{\mu(\mathbf{X})} \geq \delta$ almost surely, and

$$D(\mathbf{X}) = \frac{E[|V - \mu(\mathbf{X})| \mid \mathbf{X}]}{2\mu(\mathbf{X})} = \frac{E[|\epsilon| \mid \mathbf{X}]}{2\mu(\mathbf{X})} = \frac{E[|\epsilon|]}{2\mu(\mathbf{X})},$$

where the last equality follows from the assumption on ϵ . From Lemma 2.4.1, we have that

$$\begin{aligned}
\mathcal{R}_{XP} &= E [\mathcal{R}_{SP}(F_{V|\mathbf{X}}, c)] \\
&\geq E [\mathcal{R}_{PP}(F_{V|\mathbf{X}}, c) \cdot \alpha(S(\mathbf{X}), M(\mathbf{X}), D(\mathbf{X}))] && \text{(Theorem 2.3.1)} \\
&= E \left[(\mu(\mathbf{X}) - c) \cdot \alpha \left(\frac{S(\mathbf{X}) + M(\mathbf{X}) - 1}{M(\mathbf{X})}, 1, \frac{D(\mathbf{X})}{M(\mathbf{X})} \right) \right] && \text{(Lemma 2.3.1)} \\
&= E \left[(\mu(\mathbf{X}) - c) \cdot \alpha \left(\frac{S(\mathbf{X}) + M(\mathbf{X}) - 1}{M(\mathbf{X})}, 1, \frac{E[|\epsilon|]}{2(\mu(\mathbf{X}) - c)} \right) \right] \\
&\geq E \left[(\mu(\mathbf{X}) - c) \cdot \alpha \left(\frac{S + \delta - 1}{\delta}, 1, \frac{E[|\epsilon|]}{2(\mu(\mathbf{X}) - c)} \right) \right]. && \text{(B.18)}
\end{aligned}$$

The last inequality follows because $\alpha(S, M, D)$ is non-increasing in S , $S(\mathbf{X}) \leq S$, and the function $M \mapsto \frac{S+M-1}{M}$ is decreasing in M for $S > 1$. Since $M(\mathbf{X}) \geq \delta$, the bound follows.

By Lemma 2.3.4, $D \mapsto \alpha(\frac{S+\delta-1}{\delta}, 1, D)$ is convex. Hence, we claim that $y \mapsto y\alpha(\frac{S+\delta-1}{\delta}, 1, \frac{E[|\epsilon|]}{2y})$ is also convex whenever $y \geq 0$. Indeed, the second derivative of this function is

$$\frac{(E[|\epsilon|])^2 \cdot \frac{\partial^2}{\partial D^2} \alpha \left(\frac{S+\delta-1}{\delta}, 1, \frac{E[|\epsilon|]}{2y} \right)}{4y^3},$$

which is non-negative since $\frac{\partial^2}{\partial D^2} \alpha(\frac{S+\delta-1}{\delta}, 1, D) \geq 0$ for all D . Letting $y \rightarrow \mu(\mathbf{X}) - c$, we recognize that the right-hand side of Eq. (B.18) is an expectation of a convex function of $\mu(\mathbf{X}) - c$, and hence, by Jensen's inequality,

$$\mathcal{R}_{XP} \geq (\mu - c) \cdot \alpha \left(\frac{S + \delta - 1}{\delta}, 1, \frac{E[|\epsilon|]}{2(\mu - c)} \right) = (\mu - c) \alpha \left(\frac{M(S - 1) + \delta}{\delta}, M, \frac{E[|\epsilon|]}{2\mu} \right),$$

where the equality follows from Lemma 2.3.1. This proves the result. \square

Proof of Lemma 2.4.3. By inspection, $\mathcal{R}_{PP} = \frac{t}{2}$. To compute \mathcal{R}_{kP} , consider an optimal segmentation s_0, \dots, s_{k+1} with corresponding prices p_1, \dots, p_k . (Recall $s_0 = c = 0$, $s_{k+1} = t$, and $p_i \in [s_i, s_{i+1}), i \geq 1$.) By Lemma 2.4.2, $s_i = p_i$ for $i = 1, \dots, k$.

Now, on segment $[s_i, s_{i+1})$, the conditional distribution of V is uniform, so the personalized pricing strategy earns profit $\frac{s_{i+1}+s_i}{2} \cdot \frac{s_{i+1}-s_i}{t}$ for all i , since only $\frac{s_{i+1}-s_i}{t}$ fraction of the market is in this interval. By contrast, for $i = 1, \dots, k-1$, the k -market segmentation strategy earns revenue $s_i \frac{s_{i+1}-s_i}{t}$ since $p_i = s_i$ and thus, all customers in the segment buy. The difference in revenue between the two strategies is then

$$\begin{aligned}\mathcal{R}_{PP}(F, 0) - \mathcal{R}_{kP}(F, 0, \mathbf{s}, \mathbf{p}) &= \frac{s_1^2}{2t} + \sum_{i=1}^k \frac{s_{i+1} + s_i}{2} \cdot \frac{s_{i+1} - s_i}{t} - s_i \frac{s_{i+1} - s_i}{t} \\ &= \frac{s_1^2}{2t} + \frac{1}{2t} \sum_{i=1}^k (s_{i+1} - s_i)^2 = \frac{1}{2t} \sum_{i=0}^k (s_{i+1} - s_i)^2.\end{aligned}$$

The segmentation which maximizes $\mathcal{R}_{kP}(F, 0, \mathbf{s}, \mathbf{p})$ also minimizes this difference. By inspection, for a fixed s_1 , the optimal segmentation is equispaced, i.e., $s_i = s_{i-1} + \frac{t}{k+1}$ for $i = 1, \dots, k$. Thus $\mathcal{R}_{kP} = \sum_{i=1}^k \frac{it}{k+1} \frac{1}{k+1} = \frac{t}{2} \frac{k}{k+1}$. Consequently, $\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}} = 1 + 1/k$. \square

Proof of Theorem 2.4.2. We prove the second part of the theorem first.

(b) We first consider the case where $V \sim F$ has $\mu = 1$ and $c = 0$. Consider a partition $\delta = s_0 < s_1 < \dots < s_k < s_{k+1} = S$. Let $V_i \sim F_i$ denote the random variable V conditional on the event $V \in [s_i, s_{i+1})$, i.e., $\Pr(V_i \leq t) \equiv \Pr\left(V \leq t \mid s_i \leq V \leq s_{i+1}\right)$. Further, let $q_i = \Pr(s_i \leq V \leq s_{i+1})$ be the market share of the i -th segment, S_i be the scale of V_i and $\mathcal{R}_{SP}(F_i, 0) = \max_p p \overline{F_i}(p)$.

From Eq. (2.5) in the proof of Lemma 2.3.2 we have that for any $\gamma \leq \mathcal{R}_{SP}(F_i, 0)$,

$$\frac{\mathcal{R}_{PP}(F_i, 0)}{\mathcal{R}_{SP}(F_i, 0)} \leq 1 + \log \left(\frac{E[V_i] S_i}{\gamma} \right) \quad (\text{B.19})$$

We will apply Eq. (B.19) to each F_i with the trivial lower bound s_i . Notice that since $V_i \leq s_{i+1}$ almost surely, $S_i E[V_i] \leq x_i$.

Then, since $\mu = 1$,

$$\begin{aligned}
1 &= \sum_{i=0}^k q_i E[V_i] = \sum_{i=0}^k q_i \mathcal{R}_{PP}(F_i, 0) \\
&\leq \sum_{i=0}^k q_i \left(1 + \log \left(\frac{s_{i+1}}{s_i} \right) \right) \mathcal{R}_{SP}(F_i, 0) && (\text{since } S_i E[V_i] \leq s_{i+1}) \\
&\leq \left(\sum_{i=0}^k q_i \mathcal{R}_{SP}(F_i, 0) \right) \max_{i=0, \dots, k} \left(1 + \log \left(\frac{s_{i+1}}{s_i} \right) \right) \\
&\leq \mathcal{R}_{kP}(F, 0) \max_{i=0, \dots, k} \left(1 + \log \left(\frac{s_{i+1}}{s_i} \right) \right)
\end{aligned}$$

where the last line follows because partitioning at the s_i and offering prices $p^i \in \arg \max_{p \geq s_i} p \overline{F}_i(p)$ is a feasible segmentation policy. We minimize this last bound by setting $s_i = (\delta)^{\frac{k-i}{k}} (S)^{\frac{i}{k}}$ which implies $\frac{s_{i+1}}{s_i} = \left(\frac{S}{\delta}\right)^{1/k}$. This choice of s_i yields

$$\mathcal{R}_{PP}(F, 0) \leq \mathcal{R}_{kP}(F, 0) \left(1 + \frac{\log \left(\frac{S}{\delta} \right)}{k} \right). \tag{B.20}$$

For the general case note that for the transformation in Lemma 2.3.1, one can prove that $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kP}(F, c)} = \frac{\mathcal{R}_{PP}(F_c, 0)}{\mathcal{R}_{kP}(F_c, 0)}$ by considering each segment separately and applying an argument analogous to Lemma 2.3.1. Thus, given an F with arbitrary mean and $c > 0$, first transform to F_c and apply the above result. Substituting the original parameters proves the second part of theorem.

We now use the previous result to prove the first part of the theorem.

(a) First consider the case where $\mu = 1$ and $c = 0$. We prove the bound by separating the distribution into a small lower component near 0 and an upper component. We bound the lower component of F in terms of D and bound the upper component by applying (a). Fix some $\Delta > 1$ which we shall select later. To bound the lower

tail, note

$$\begin{aligned} D &= E[(1 - V)^+] \geq E\left[(1 - V)^+ \mathbb{I}\left(V \leq \frac{1}{\Delta}\right)\right] \\ &\geq E\left[\left(1 - \frac{1}{\Delta}\right) \mathbb{I}\left(V \leq \frac{1}{\Delta}\right)\right] = \left(1 - \frac{1}{\Delta}\right) \Pr\left(V \leq \frac{1}{\Delta}\right). \end{aligned}$$

Rearranging yields $\Pr(V \leq 1/\Delta) \leq \frac{\Delta D}{\Delta - 1}$. This further implies that

$$E\left[V \mathbb{I}\left(V \leq \frac{1}{\Delta}\right)\right] \leq \frac{1}{\Delta} \Pr(V \leq 1/\Delta) \leq \frac{D}{\Delta - 1}. \quad (\text{B.21})$$

Now by splitting the expectation,

$$\begin{aligned} 1 = \mathcal{R}_{PP}(F, 0) &= E\left[V \mathbb{I}\left(V \leq \frac{1}{\Delta}\right)\right] + E\left[V \mathbb{I}\left(V \geq \frac{1}{\Delta}\right)\right] \\ &\leq \frac{D}{\Delta - 1} + E\left[V \mathbb{I}\left(V \geq \frac{1}{\Delta}\right)\right] \quad (\text{using Eq. (B.21)}) \\ &= \frac{D}{\Delta - 1} + E[V_\Delta] \Pr(V \geq 1/\Delta) \end{aligned}$$

where V_Δ is the conditional distribution of V given that $V \geq 1/\Delta$, i.e., $\Pr(V_\Delta \geq t) = \Pr(V \geq t | V \geq 1/\Delta)$.

Note that $E[V_\Delta] \geq 1$ and that V_Δ has scale $S/E[V_\Delta] \leq S$. Most importantly, $\Pr(V_\Delta \geq \frac{1}{\Delta}) = 1$, so that we can apply part (b) to upper bound the expectation yielding

$$\mathcal{R}_{PP}(F, 0) \leq \frac{D}{\Delta - 1} + \left(1 + \frac{\log(S\Delta)}{k}\right) R_{kP}(F_\Delta, 0) \Pr(V \geq 1/\Delta).$$

Finally, letting \setminus^* be the optimal segmentation for $\mathcal{R}_{kP}(F_\Delta, 0)$ and \mathbf{p}^* be the corre-

sponding prices. Define the function

$$r(v) = \begin{cases} p_i^* & \text{if } v \in [s_i^*, s_{i+1}^*) \text{ and } v \geq p_i^*, \quad i = 1, \dots, k \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned} \mathcal{R}_{kP}(F, 0) &\geq \mathcal{R}_{k,P}(F, 0, \setminus^*, \mathbf{p}^*) \\ &= E[r(V)] \\ &= E \left[r(V) \mid V \leq \frac{1}{\Delta} \right] \Pr \left(V \leq \frac{1}{\Delta} \right) + E \left[r(V) \mid V \geq \frac{1}{\Delta} \right] \Pr \left(V \geq \frac{1}{\Delta} \right) \\ &= E \left[r(V) \mid V \leq \frac{1}{\Delta} \right] \Pr \left(V \leq \frac{1}{\Delta} \right) + R_{kP}(V_\Delta) \Pr \left(V \geq \frac{1}{\Delta} \right) \\ &\geq R_{kP}(V_\Delta) \Pr \left(V \geq \frac{1}{\Delta} \right) \end{aligned}$$

Plugging in above yields,

$$\mathcal{R}_{PP}(F, 0) \leq \frac{D}{\Delta - 1} + \left(1 + \frac{\log(S\Delta)}{k} \right) R_{kP}(F, 0).$$

Letting $\Delta = 1 + D(k + 1)$ and rearranging yields

$$\frac{\mathcal{R}_{PP}}{\mathcal{R}_{kP}} \leq \frac{1 + \frac{\log(S\Delta)}{k}}{1 - \frac{D}{\Delta - 1}} \leq 1 + \frac{1}{k} + \frac{\log(S + SD(k + 1))}{k} = 1 + O\left(\frac{\log(k)}{k}\right).$$

For a general $c > 0$ and $\mu \neq 1$, apply the transformation of Lemma 2.3.1. As in the previous part, note that $\frac{\mathcal{R}_{PP}(F, c)}{\mathcal{R}_{kP}(F, c)} = \frac{\mathcal{R}_{PP}(F_c, 0)}{\mathcal{R}_{kP}(F_c, 0)}$. Apply the result of the previous part and then make the appropriate substitutions. \square

Proof of Theorem 2.4.3. We shall prove that

$$\mathcal{R}_{kP}(F_{\mu(\mathbf{x})}, c) - \mathcal{R}_{kXP}(F_{\mathbf{x}V}, c) \leq \mathcal{R}_{PP}(F_{\mu(\mathbf{x})}, c) - \mathcal{R}_{XP}(F_{\mathbf{x}V}, c). \quad (\text{B.22})$$

Note that $\mathcal{R}_{PP}(F_{\mu(\mathbf{x})}, c) = \mathcal{R}_{PP}(F, c)$ so that Eq. (B.22) implies the two inequalities above by rearranging. In fact, we will prove that the portion of profits earned by each strategy for a fixed context $\mathbf{x} \in \mathcal{X}$ satisfies Eq. (B.22).

To that end, let $\{s_i\}_{i=0}^k \in \mathbb{R}^{k+1}$ be an optimal k -market segmentation for $\mathcal{R}_{kP}(F_{\mu(\mathbf{x})})$ of the form described in Lemma 2.4.2 where $s_0 = c$ and $s_{k+1} = \infty$. Partition the feature space, $\mathcal{X}_i := \{\mathbf{x} \in \mathcal{X} \mid \mu(\mathbf{x}) \in [s_i, s_{i+1})\}$ for $i \in 0, 1, \dots, k$. Let $p(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$ denote the optimal pricing function for $\mathcal{R}_{XP}(F_{\mathbf{x}V}, c)$. Finally, let $\mathbf{x} \in \mathcal{X}_i$ be some fixed realization of \mathbf{X} for some $i = 0, \dots, k$.

When $\mathbf{X} = \mathbf{x}$, the XP strategy earns

$$\text{x-Contribution to XP} = (p(\mathbf{x}) - c)\mathbb{P}\{\mu(\mathbf{x}) + \epsilon \geq p(\mathbf{x})\},$$

where we have used independence to drop the conditioning on $\mathbf{X} = \mathbf{x}$.

Similarly, when $\mathbf{X} = \mathbf{x}$, the kP strategy earns

$$\text{x-Contribution to kP} = (s_i - c)$$

on $F_{\mu(\mathbf{x})}$, since by Lemma 2.4.2 all customers in the segment buy at price s_i .

Next, we lower bound the **x-Contribution to kXP** by considering a feasible feature-based segmentation strategy. Let $\mathbf{x}_i \in \arg \min_{\mathbf{x} \in \mathcal{X}_i} \mu(\mathbf{X})$, and consider offering price $p(\mathbf{x}_i)$ to every customer in \mathcal{X}_i . When $\mathbf{X} = \mathbf{x}$, this strategy earns at most

$$(p(\mathbf{x}_i) - c)\mathbb{P}\{\mu(\mathbf{x}) + \epsilon \geq p(\mathbf{x}_i)\} \geq (p(\mathbf{x}_i) - c)\mathbb{P}\{\mu(\mathbf{x}_i) + \epsilon \geq p(\mathbf{x}_i)\},$$

since $\mu(\mathbf{x}) \geq \mu(\mathbf{x}_i)$ for all $\mathbf{x} \in \mathcal{X}_i$. We again use independence to drop conditioning on $\mathbf{X} = \mathbf{x}$. On the other hand, by Lemma 2.4.1, $p(\mathbf{x}_i)$ is the optimal single price

when $\mathbf{X} = \mathbf{x}_i$, so that pricing at $p(\mathbf{x}) + \mu(\mathbf{x}_i) - \mu(\mathbf{x})$ must earn less profit, i.e.,

$$\begin{aligned}
& (p(\mathbf{x}_i) - c) \mathbb{P} \{ \mu(\mathbf{x}_i) + \epsilon \geq p(\mathbf{x}_i) \} \\
& \geq (p(\mathbf{x}) + \mu(\mathbf{x}_i) - \mu(\mathbf{x}) - c) \mathbb{P} \{ \mu(\mathbf{x}_i) + \epsilon \geq p(\mathbf{x}) + \mu(\mathbf{x}_i) - \mu(\mathbf{x}) \} \\
& = (p(\mathbf{x}) + \mu(\mathbf{x}_i) - \mu(\mathbf{x}) - c) \mathbb{P} \{ \mu(\mathbf{x}) + \epsilon \geq p(\mathbf{x}) \},
\end{aligned}$$

where now we use independence to drop the conditioning $\mathbf{X} = \mathbf{x}_i$ throughout. Combining these two inequalities shows

$$\text{x-Contribution to kXP} \geq (p(\mathbf{x}) + \mu(\mathbf{x}_i) - \mu(\mathbf{x}) - c) \mathbb{P} \{ \mu(\mathbf{x}) + \epsilon \geq p(\mathbf{x}) \}.$$

Now combine these contributions as in Eq. (B.22),

$$\begin{aligned}
& \text{x-Contribution to XP} + \text{x-Contribution to kP} - \text{x-Contribution to kXP} \\
& \leq (s_i - c) + (\mu(\mathbf{x}) - \mu(\mathbf{x}_i)) \mathbb{P} \{ \mu(\mathbf{x}) + \epsilon \geq p(\mathbf{x}) \} \\
& \leq (s_i - c) + (\mu(\mathbf{x}) - \mu(\mathbf{x}_i)) \\
& \leq \mu(\mathbf{x}) - c,
\end{aligned}$$

where the last line follows because $\mu(\mathbf{x}_i) \in [s_i, s_{i+1})$. Recall \mathbf{x} was chosen arbitrarily. Averaging this inequality over realizations of \mathbf{X} yields Eq. (B.22) to complete the proof. \square

Omitted Proofs from Section 2.5.

Proof of Theorem 2.5.1. Without loss of generality, suppose $c = 0$. Now recall that the revenue of a feature based strategy $\mathcal{R}_{XP}(F, 0)$ equals the expectation of single

price strategies over every realization of \mathbf{X} i.e.,

$$\begin{aligned}\mathcal{R}_{XP}(F, 0) &= E[\mathcal{R}_{SP}(F_{V|\mathbf{X}}, 0)] \\ &= \int_c^\infty \mathcal{R}_{SP}(F_{x+\epsilon}, 0) f_{\mu(\mathbf{X})=x} dx = \int_c^\infty x \frac{\mathcal{R}_{SP}(F_{x+\epsilon}, 0)}{x} f_{\mu(\mathbf{X})=x} dx \quad (\text{B.23})\end{aligned}$$

where as in Theorem 2.4.3, the independence of \mathbf{X} and ϵ allows us to drop of the conditioning $\mu(\mathbf{X}) = \mathbf{x}$. We will prove bounds on Eq. (B.23) by applying the bounds from Theorems 2.3.1 and 2.3.2 to each realization of \mathbf{X} , $\frac{\mathcal{R}_{SP}(F_{x+\epsilon})}{x}$. Note the coefficient of deviation of the distribution $x + \epsilon$ is $\frac{E[|\epsilon|]}{2x}$, and recall by Theorem 2.3.1(a), $\frac{\mathcal{R}_{SP}(F_{x+\epsilon})}{x} \geq \frac{1 - \frac{E[|\epsilon|]}{2x}}{-W_{-1}\left(\frac{\frac{E[|\epsilon|]}{2x} - 1}{e}\right)}$. Since $x \frac{1 - \frac{E[|\epsilon|]}{2x}}{-W_{-1}\left(\frac{\frac{E[|\epsilon|]}{2x} - 1}{e}\right)}$ is convex in x on the open interval $[c, \infty)$, applying Jensen's inequality to the right hand side of Eq. (B.23) yields the desired lower bound.

To derive an upper bound, first note $x + \epsilon$ is unimodal and left-skew for any x , thus the conditions of Theorem 2.3.2 are satisfied for every realization of \mathbf{X} and may be applied in a similar fashion as before. Unfortunately the resulting integral is convex in x . To circumvent this obstacle, it can be easily checked that

$$\frac{2-D}{2} \geq \begin{cases} 1-D & \text{if } D \leq 1/3, \\ \frac{(1+D)^2}{8D} & \text{if } D \geq 1/3 \end{cases}.$$

and further $\frac{2-D}{2}$ is the minimal concave (linear in fact) upper bound on the right hand side of the reciprocal of Theorem 2.3.2. Applying this looser bound to Eq. (B.23) yields,

$$\int_c^\infty x \frac{\mathcal{R}_{SP}(F_{x+\epsilon}, 0)}{x} f_{\mu(\mathbf{X})=x} dx \leq \int_c^\infty x \left(\frac{2 - \frac{E[|\epsilon|]}{2x}}{2} \right) f_{\mu(\mathbf{X})=x} dx = \mu \frac{2 - D_{\mu\epsilon}}{2}.$$

Dividing through by $\mathcal{R}_{SP}(F, 0)$ gives the result. \square

Proof of Theorem 2.5.2. We first consider the case when $c = 0$ and $\mu = 1$ so that $M = 1$. Our strategy will be to compute lower bound, z^* , on $\mathcal{R}_{SP}/\mathcal{R}_{PP}$ so that $1/z^*$ is an upper bound on $\mathcal{R}_{PP}/\mathcal{R}_{SP}$. Since (2.11) provides a tight upper bound on $\mathcal{R}_{PP}/\mathcal{R}_{SP}$, it follows immediately that the optimization problem

$$\begin{aligned}
& \inf_{y, d\mathbb{P}_v} \quad y \\
& \text{s.t.} \quad \int_0^S d\mathbb{P}_v = 1 \\
& \quad \int_0^S v d\mathbb{P}_v = 1, \quad \int_0^S f(v) d\mathbb{P}_v = 0, \\
& \quad y \geq p \int_0^S \mathbb{I}(v \geq p) d\mathbb{P}_v, \quad \forall p \in [0, S]. \\
& \quad d\mathbb{P}_v \geq 0 \quad \forall v \in [0, S]
\end{aligned} \tag{B.24}$$

is a tight lower bound on $\mathcal{R}_{SP}/\mathcal{R}_{PP}$. Following [95], the dual to this optimization problem is

$$\begin{aligned}
& \sup_{\theta, \lambda, d\mathcal{Q}_p} \quad \theta + \lambda_1 \\
& \text{s.t.} \quad \int_0^S d\mathcal{Q}_p = 1 \\
& \quad \theta + \lambda_1 v + \lambda_2 f(v) - \int_0^S p \mathcal{I}(v \geq p) d\mathcal{Q}_p \leq 0, \quad \forall v \in [0, S] \\
& \quad d\mathcal{Q}_p \geq 0 \quad \forall p \in [0, S].
\end{aligned} \tag{B.25}$$

By weak-duality, any feasible solution to problem (B.25) yields a valid lower bound to (B.24). To form such a feasible solution to (B.25), we constrain \mathcal{Q}_p to be supported only on $\{p_0, \dots, p_N\}$ and denote the corresponding point masses as Q_0, Q_1, \dots, Q_N .

Then, the value of (B.25) is at least

$$\begin{aligned}
z^* &:= \max_{\theta, \lambda, \mathbf{Q}} \quad \theta + \lambda_1 \\
\text{s.t.} \quad & \sum_{j=0}^N Q_j = 1, \quad Q_j \geq 0 \quad \forall j = 0, \dots, N \quad (\text{B.26}) \\
& \theta + \lambda_1 v + \lambda_2 f(v) - \sum_{j=0}^N p_j \mathcal{I}(v \geq p_j) Q_j \leq 0, \quad \forall v \in [0, S] .
\end{aligned}$$

Notice that the sum of indicators in the second set of constraints of Eq. (B.26) is constant over $[p_{k-1}, p_k)$. Thus, we can rewrite this constraint of Eq. (B.26) as $N + 1$ separate sets of constraints:

$$\begin{aligned}
\theta + \lambda_1 v + \lambda_2 f(v) &\leq \sum_{j=0}^{k-1} p_j Q_j, \quad \forall v \in [p_{k-1}, p_k), \quad k = 1, \dots, N, \\
\theta + \lambda_1 S + \lambda_2 f(S) &\leq \sum_{j=0}^N p_j Q_j .
\end{aligned}$$

Replacing the second family of constraints of Eq. (B.26) with these $N + 1$ sets of constraints completes the proof when $\mu = 1$ and $c = 0$.

In the general case, by Lemma 2.3.1, the value of personalization for a random variable V with scale S , margin M and mean μ equals the value of personalization for the random variable $V_c = \frac{V-c}{\mu-c}$, which has scale $S_c = \frac{S+M-1}{M}$, margin 1, and mean 1. Furthermore, if $E[f(V)] = 0$, then, $E[f_c(V_c)] = 0$ where $f_c(t) = f(t(\mu - c) + c)$. Thus, we can apply the above result to V_c (with shifted parameters) and $f_c(\cdot)$ with the same discretization points p_j , $j = 0, \dots, N$, yielding the following lower-bound

on $\mathcal{R}_{SP}/\mathcal{R}_{PP}$:

$$\begin{aligned}
z^* &:= \max_{\theta, \lambda, \mathbf{Q}} \theta + \lambda_1 \\
\text{s.t.} \quad & \sum_{j=0}^N Q_j = 1, \quad Q_j \geq 0 \quad \forall j = 0, \dots, N \\
& \theta + \lambda_1 v + \lambda_2 f_c(v) \leq \sum_{j=0}^{k-1} p_j Q_j, \quad \forall v \in [p_{k-1}, p_k], \quad k = 1, \dots, N, \\
& \theta + \lambda_1 S_c + \lambda_2 f_c(S_c) \leq \sum_{j=0}^N p_j Q_j.
\end{aligned} \tag{B.27}$$

Note that $f_c(v) = f(v(\mu - c) + c) = f(v\mu M + \mu(1 - M))$ and $f_c(S_c) = f(S\mu)$ to complete the proof. \square

Proof. Proof of Lemma 2.5.1. Let p^* be such that $\sup_{F \in \mathcal{F}} \mathcal{R}_{SP}(F) = \sup_{F \in \mathcal{F}} \mathcal{R}_{SP}(F, p^*)$ and let i be such that $(1 + \delta)^i - 1 \leq p^* \leq (1 + \delta)^{i+1} - 1$. Let F^*, F_i, F_{i+1} be the respective tight distributions and note that each distribution is in \mathcal{F} . Then,

$$\begin{aligned}
p^* \overline{F^*}(p^*) &\leq p_{i+1} \overline{F^*}(p^*) \leq p_{i+1} \overline{F^*}(p_i) \leq p_{i+1} \overline{F_i}(p_i) \\
&= (1 + \delta) p_i \overline{F_i}(p_i) \leq (1 + \delta) \sup_{F \in \mathcal{F}} \max_{p \in \Delta} \mathcal{R}_{SP}(F, p)
\end{aligned}$$

Where the third inequality follows by the optimality of $\overline{F_i}$ for (P_i) and all other inequalities follow from their respective definitions. \square

B.3 A Dynamic Programming Algorithm for Computing (kP)

In this section, we describe an efficient dynamic programming algorithm for computing the optimal k -market segmentation when the valuation distribution is known

precisely and discretely supported on n values. One should compare this algorithm to the distribution-agnostic procedure given in Theorem 2.4.2 when the valuation distribution is not known precisely.

Structurally, computing the optimal k -market segmentation is extremely similar to the *1D Clustering* problem for which dynamic programming approaches have been employed (see [64] for a modern overview). Formally, suppose V is supported on n values $\{v_i\}_{i=1}^n$, occurring with probabilities $\{q_i\}_{i=1}^n$. Without loss of generality suppose the values are indexed from low to high, i.e., $v_i \leq v_{i+1}$ for all i . By an argument identical to Lemma 2.4.2 the optimal segmentation $\{s_i\}_{i=0}^k$ is contained in the support of V , we wish to find $\{s_i\}_{i=0}^k \subset \{v_i\}_{i=1}^n$ that maximizes

$$\sum_{i=1}^k s_i \Pr(V \in [s_i, s_{i+1})).$$

We give a dynamic programming solution that uses time $O(kn^2)$. Define $D[m, j]$ as the optimal j -market segmentation that considers the m lowest points $\{(v_i, q_i)\}_{i=1}^m$, our goal is to compute $D[n, k]$. Our algorithm depends on the following observation: consider the optimal k -market segmentation and suppose $[v_{i_k}, v_n]$ defines the k^{th} segment. If one considers the market without the customers in the k^{th} segment, the remaining $k - 1$ segments must be an optimal $(k - 1)$ -market segmentation on $\{(v_i, q_i)\}_{i=1}^{i_k-1}$. Formally we express this observation as the following recursion,

$$D[m, j] = \max_{l \in [m-1]} D[l, j-1] + v_{l+1} \sum_{i=l+1}^m q_i, \quad (\text{B.28})$$

which states that the optimal j -market segmentation on the lowest m valuations, is equal to some optimal $(j - 1)$ -segmentation on a smaller market, plus the value of the j^{th} segment. Using Eq. (B.28) we may populate a table of size kn , starting at $D[0, 0] = 0$, and computing column-wise. Each computation of $D[m, j]$ requires $O(n)$ operations, thus the table may be populated in $O(kn^2)$ time.

Chapter C

Appendix to Chapter 3

C.1 Omitted Proofs from Chapter 3

Omitted Proofs from Section 3.2

Proof of Theorem 3.2.1. We will consider the two cases separately.

(a) Unique Box. Consider a customer facing a catalog of N items. Let $X_k = \sum_{i=1}^k (V_i - p)$, which is the net utility of buying k unique boxes. Note the customer will buy the first unique box if $X_N \geq 0$, and continue to purchase until the mean valuation for the remaining items is less than p . Specifically, a myopic customer stops purchasing at the first t such that $\frac{\sum_{i \notin S_t} (V_i - p)}{N-t} < 0$, which implies that

$$\sum_{i \in S_t} (V_i - p) = X_N - \sum_{i \notin S_t} (V_i - p) \geq X_N.$$

Hence, the net utility of myopic strategy is at least X_N . On the other hand, consider $\{V_i\}_{i=1}^N$ as valuations for the goods in the order in which they are purchased i.e. the i^{th} loot box purchase yields an item valued at V_i . Then an upper bound on the maximum possible net utility of any purchasing strategy is the utility of the clairvoyant strategy that stops when customer utility is maximized: $M_N = \max_{k \in [N]} \sum_{i=1}^k (V_i - p)$.

For a random customer, since V_i are i.i.d, M_N is equivalent to the maximum deviation from 0 of the random walk of partial sums, $\{X_k\}_{k=1}^N$. By Theorem 2.12.1 in [65], $\lim_{N \rightarrow \infty} M_N/N$ converges to $\max(0, \mu - p)$ almost surely. Further by the strong

law of large numbers, $\lim_{N \rightarrow \infty} X_N/N$ converges to $\mu - p$ almost surely. Let U_N^{OPT} and U_N^M be the net utility of the optimal strategy and the myopic strategy, respectively. Then we have $\max(0, X_N) \leq U_N^M \leq U_N^{OPT} \leq M_N$. Dividing by N yields:

$$\frac{U_N^M}{N} \geq \frac{\max(0, X_N)}{N} \xrightarrow{a.s.} \max(0, \mu - p) \quad \text{and} \quad \frac{U_N^M}{N} \leq \frac{U_N^{OPT}}{N} \leq \frac{M_N}{N} \xrightarrow{a.s.} \max(0, \mu - p)$$

which together imply that the normalized net utility of both the optimal strategy and the myopic strategy converge to $\max(0, \mu - p)$ almost surely.

(b) Traditional Box. As in (a) consider a customer facing a catalog of N items, with valuations $\{V_i\}_{i=1}^N$. If a customer purchased t boxes and their valuation for another traditional box exceeds the price, $\frac{1}{N} \sum_{j \in [N] - S_t} V_j \geq p$, then clearly purchasing the next box gives positive expected utility and will be undertaken by an optimal policy. Suppose instead that after t purchases $\frac{1}{N} \sum_{j \in [N] - S_t} V_j < p$. We will show that in this case a purchase will never increase a customers utility in expectation which implies the myopic policy is optimal.

To see this, consider the following lottery: with the probability $\frac{1}{N}$ the customer will receive V_j utility for each $j \notin S_t$, and with probability $\frac{N - |S_t|}{N}$ they will receive nothing. Clearly after t purchases this lottery is equivalent to a traditional loot box, and for every loot box purchase after t , the customer will prefer this lottery since the lottery eliminates the replacement effect from items acquired after t . However, even for this strictly better lottery, buying any amount has negative utility. Hence when customers reach a traditional box for which $\frac{1}{N} \sum_{j \in [N] - S_t} V_j < p$, it is optimal to stop purchasing. \square

Omitted Proofs from Section 3.3

Proof of Item a). Consider the extended sum utility random walk from the proof of Theorem 3.2.1(a), $\{X_j\}_{j=0}^\infty$ where $X_j := \sum_{i=1}^j (V_i - p)$, $X_0 = 0$, where valuations are indexed so that the i^{th} item a customer receives is valued at V_i . Recall a myopic

customer will purchase loot boxes until their expected utility is negative i.e. until the first time t such that $X_t > X_N$.

To prove the theorem, we will construct a sequence of prices $p_N = \mu - \epsilon_N$, where $\lim_{N \rightarrow \infty} \epsilon_N = 0$. Then X_t is always a random walk with positive drift, and we show that the fraction of unique boxes purchased under these prices tends to 1. Fix an N , a corresponding price $p_N = \mu - \epsilon_N < \mu$, and some $k_N \in (0, 1)$. Let s_N be the selling volume, i.e., the first time t such that $X_t > X_N$, and τ_N be the first passage time of $\{X_i\}$ to the line $(1 - k_N)N\epsilon_N$. Unfortunately s_N is not a stopping time making it difficult to characterize, however τ_N is, and we will use τ_N to approximate s_N for our desired result. Note that for the sample paths such that $X_N \geq (1 - k_N)N\epsilon_N$, $s_N \geq \tau_N$ because it must hit $(1 - k_N)N\epsilon_N$ before hitting X_N . Hence we have,

$$\begin{aligned} \mathbb{E}[s_N] &\geq \mathbb{E}[s_N \mathcal{I}_{X_N \geq (1-k_N)N\epsilon_N}] \\ &\geq \mathbb{E}[\tau_N \mathcal{I}_{X_N \geq (1-k_N)N\epsilon_N}] \\ &= \mathbb{E}[\tau_N] - \mathbb{E}[\tau_N \mathcal{I}_{X_N \in [0, (1-k_N)N\epsilon_N)}] - \mathbb{E}[\tau_N \mathcal{I}_{X_N < 0}]. \end{aligned} \quad (\text{C.1})$$

We will lower bound Eq. (C.1) term by term. First since τ_N is a stopping time, we can compute $\mathbb{E}[\tau_N]$ by rearranging Wald's equation $\mathbb{E}[X_{\tau_N}] = \mathbb{E}[\tau_N]\epsilon_N$ yielding,

$$\mathbb{E}[\tau_N] = \frac{\mathbb{E}[X_{\tau_N}]}{\epsilon_N} \geq \frac{(1 - k_N)N\epsilon_N}{\epsilon_N} = (1 - k_N)N. \quad (\text{C.2})$$

For the second term in (C.1), i.e., the case that $X_N \in [0, (1 - k_N)N\epsilon_N)$, the most conservative case for the seller is when $X_N = 0$ and the walk does not hit $(1 - k_N)N\epsilon_N$ before N . In this case it takes another $(1 - k_N)N$ steps in expectation to reach $(1 - k_N)N\epsilon_N$. Thus,

$$\mathbb{E}[\tau_N | X_N \in [0, (1 - k_N)N\epsilon_N)] \leq N + (1 - k_N)N. \quad (\text{C.3})$$

The probability that $X_N \in [0, (1 - k_N)N\epsilon_N]$ can be upper bounded using Chebyshev's Inequality,

$$\mathbb{P}(X_N \in [0, (1 - k_N)N\epsilon_N]) \leq \mathbb{P}(X_N < (1 - k_N)N\epsilon_N) \leq \frac{\sigma^2}{k_N^2 \epsilon_N^2 N}. \quad (\text{C.4})$$

Combining (C.3) and (C.4), we have

$$\mathbb{E}[\tau_N \mathcal{I}_{X_N \in [0, (1 - k_N)N\epsilon_N]}] \leq (N + (1 - k_N)N) \frac{\sigma^2}{k_N^2 \epsilon_N^2 N} = \frac{(2 - k_N)\sigma^2}{k_N^2 \epsilon_N^2 N}. \quad (\text{C.5})$$

For the third term in (C.1), i.e., the case that $X_N < 0$, since $V_i \geq 0$, a simple lower bound on X_N is $-Np_N$. Supposing $X_N = -Np_N$, in expectation it takes another $(Np_N + (1 - k_N)N\epsilon_N)/\epsilon_N$ steps from N for the random walk to hit $(1 - k_N)N\epsilon_N$. Thus, $\mathbb{E}[\tau_N | X_N < 0] \leq N + \frac{Np_N + (1 - k_N)N\epsilon_N}{\epsilon_N}$ and, as before, the probability that $X_N < 0$ can also be upper bounded using Chebyshev's Inequality,

$$\mathbb{P}(X_N < 0) \leq \frac{\sigma^2}{\epsilon_N^2 N}.$$

Combining these two observations yields,

$$\begin{aligned} \mathbb{E}[\tau_N \mathcal{I}_{X_N < 0}] &\leq \left(N + \frac{Np_N + (1 - k_N)N\epsilon_N}{\epsilon_N} \right) \frac{\sigma^2}{\epsilon_N^2 N} \\ &= \left(1 + \frac{\mu - \epsilon_N}{\epsilon_N} + (1 - k_N) \right) \frac{\sigma^2}{\epsilon_N^2} = \left(\frac{\mu}{\epsilon_N} + 1 - k_N \right) \frac{\sigma^2}{\epsilon_N^2} \end{aligned} \quad (\text{C.6})$$

Plugging Eqs. (C.2), (C.5) and (C.6) into the right hand side of Eq. (C.1) yields,

$$\mathbb{E}[s_N] \geq N \left(1 - k_N - \frac{2\sigma^2}{k_N^2 \epsilon_N^2 N} + \frac{\sigma^2}{k_N \epsilon_N^2 N} - \frac{\sigma^2 \mu}{\epsilon_N^3 N} - \frac{\sigma^2}{\epsilon_N^2 N} + \frac{\sigma^2 k_N}{\epsilon_N^2 N} \right). \quad (\text{C.7})$$

Now we can lower bound the normalized revenue of a unique box strategy,

$$\begin{aligned}\mathcal{R}_{UB} &\geq (\mu - \epsilon_N) \frac{\mathbb{E}[s_N]}{N} \\ &\geq \mu(1 - \frac{\epsilon_N}{\mu}) \left(1 - k - \frac{2\sigma^2}{k^2\epsilon_N^2 N} + \frac{\sigma^2}{k\epsilon_N^2 N} - \frac{\sigma^2\mu}{\epsilon_N^3 N} - \frac{\sigma^2}{\epsilon_N^2 N} + \frac{\sigma^2 k}{\epsilon_N^2 N} \right)\end{aligned}\quad (\text{C.8})$$

Choosing $\epsilon_N = \mu N^{-1/5}$, and $k_N = N^{-1/5}$, we have

$$\mathcal{R}_{UB} \geq \mu(1 - N^{-1/5}) \left(1 - (1 + \frac{2\sigma^2}{\mu^2})N^{-1/5} - \frac{\sigma^2}{\mu^2}N^{-3/5} + \frac{\sigma^2}{\mu^2}N^{-4/5} \right).$$

Taking the limit of both sides gives

$$\lim_{N \rightarrow \infty} \mathcal{R}_{UB} \geq \mu.$$

Combined with the fact that $\mathcal{R}_{UB} \leq \mu$, we conclude that $\lim_{N \rightarrow \infty} \mathcal{R}_{UB} = \mu$. \square

Proof of Item b). Consider the sum of valuations random walk, $\{Y_i\}_{i=0}^\infty$, where $Y_j = \sum_{i=1}^j V_i$ and $Y_0 = 0$. For a random customer, Y_N/N is their expected valuation of the first traditional box. Each time the customer receives a new item, they move back one index on the random walk from Y_{j+1} to Y_j , the customers total valuation for all the remaining j items. Note that Y_i is strictly increasing with i , and every time the customer receives a new item, their total valuation for the remaining item decreases monotonically.

Similar to our proof of (a), we will construct a sequence of prices p_N such that $\lim_N p_N \rightarrow \frac{\mu}{e}$ and show the expected number of traditional loot boxes purchased by a customer at price p_N tends to N . Fix an N and a corresponding price p_N . Then a customer purchases until the first time t (counting down from N) such that $\frac{Y_t}{N} < p_N \iff Y_t < Np_N$. Let $\tau(p_N)$ be the first passage time of $\{Y_i\}$ to Np_N i.e., $\tau(p_N) := \min\{t : Y_t \geq Np_N\}$. Note that since Y_i is monotonic, the random walk crosses Np_N exactly once. Suppose $\tau(p_N) \leq N$, then at time $\tau(p_N)$ the number of

distinct items a customer will own is $N - \tau(p_N) - 1$, and the number of traditional loot boxes they will have purchased is the sum of $N - \tau(p_N) - 1$ independent geometric random variables, $\text{Geo}(1) + \text{Geo}(\frac{N-1}{N}) + \dots + \text{Geo}(\frac{\tau(p_N)}{N})$. The revenue under price p_N is then,

$$\begin{aligned}
\mathcal{R}_{TB}(p_N) &= \frac{1}{N} \mathbb{E}[p_N \left(\text{Geo}(1) + \text{Geo}\left(\frac{N-1}{N}\right) + \dots + \text{Geo}\left(\frac{\tau(p_N)}{N}\right) \right) \mathcal{I}_{\tau(p_N) \leq N}] \\
&= p_N \mathbb{E} \left[\frac{1}{N} \left(1 + \frac{N}{N-1} + \dots + \frac{N}{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq N} \right] \\
&= p_N \mathbb{E} \left[(\log(N) + \gamma + \zeta_N - \log(\tau(p_N)) - \gamma - \zeta_{\tau(p_N)}) \mathcal{I}_{\tau(p_N) \leq N} \right] \\
&= p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} + \zeta_N - \zeta_{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq N} \right], \tag{C.9}
\end{aligned}$$

where the third equality follows from the well known expression for the harmonic numbers, $\sum_{k=1}^n \frac{1}{k} = \log n + \gamma + \zeta_n$, with $\{\zeta_k\}$ converges to 0 from above, and γ is the EulerMascheroni constant.

First we will bound $\mathbb{E}[\tau(p_N)]$. Since $\tau(p_N)$ is the passage time, $Y_{\tau(p_N)} \geq Np$ and $Y_{\tau(p_N)-1} < Np$, it follows by the well known inspection paradox that $\mathbb{E}[X_{\tau(p_N)}] = \frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]} = \frac{\mu^2 + \sigma^2}{\mu}$. Using this fact together with Wald's equation, $\mathbb{E}[Y_{\tau(p_N)}] = \mathbb{E}[\tau(p_N)]\mu$, we have

$$\mathbb{E}[\tau(p_N)] = \frac{\mathbb{E}[Y_{\tau(p_N)}]}{\mu} \in \left[\frac{Np}{\mu}, \frac{Np}{\mu} + 1 + \frac{\sigma^2}{\mu^2} \right]. \tag{C.10}$$

Now we can construct a lower bound for $\mathcal{R}_{TB}(N, p_N)$,

$$\begin{aligned}
\mathcal{R}_{TB}(N, p_N) &= p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} + \zeta_N - \zeta_{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq N} \right] && (Eq. (C.9)) \\
&\geq p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} \right) \mathcal{I}_{\tau(p_N) \leq N} \right] && (\{\zeta_k\} \text{ monotone dec.}) \\
&\geq p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} \right) \right] && (C.11) \\
&\geq -p_N \log \mathbb{E} \left[\frac{\tau(p_N)}{N} \right] && (\text{Jensen's Inequality}) \\
&\geq -p_N \log \left(\frac{p_N}{\mu} + \frac{1 + \frac{\sigma^2}{\mu^2}}{N} \right) && (Eq. (C.10)) \quad (C.12)
\end{aligned}$$

where Eq. (C.11) follows from the fact that $-\log \frac{\tau(p_N)}{N} < 0$ when $\tau(p_N) > N$. Setting $p_N = \frac{\mu}{e}$ yields

$$\mathcal{R}_{TB}(N) \geq \frac{\mu}{e} \log \left(\frac{1}{\frac{1}{e} + \frac{1 + \frac{\sigma^2}{\mu^2}}{N}} \right) \quad (C.13)$$

which is our desired guarantee. We will now upper bound the revenue, $\mathcal{R}_{TB}(N, p_N)$.

Consider the event that $\frac{\tau(p_N)}{N} \leq (1 - \epsilon_N) \frac{p}{\mu}$ for some small ϵ_N . We can upper bound

the probability of such an event by,

$$\begin{aligned}
\mathbb{P}\left(\frac{\tau(p_N)}{N} \leq (1 - \epsilon_N) \frac{p_N}{\mu}\right) &= \mathbb{P}\left(\tau(p_N) \leq (1 - \epsilon_N) \frac{Np_N}{\mu}\right) \\
&= \mathbb{P}\left(\sum_{t=1}^{(1-\epsilon_N) \frac{Np_N}{\mu}} V_t \geq Np_N\right) \\
&= \mathbb{P}\left(\frac{\sum_{t=1}^{(1-\epsilon_N) \frac{Np_N}{\mu}} V_t}{(1 - \epsilon_N) \frac{Np_N}{\mu}} \geq \frac{1}{1 - \epsilon_N} \mu\right) \\
&\leq \mathbb{P}\left(\left|\frac{\sum_{t=1}^{(1-\epsilon_N) \frac{Np_N}{\mu}} V_t}{(1 - \epsilon_N) \frac{Np_N}{\mu}} - \mu\right| \geq \frac{\epsilon_N}{1 - \epsilon_N} \mu\right) \\
&\leq \frac{\sigma^2}{(1 - \epsilon_N) \frac{Np_N}{\mu} \frac{\epsilon_N^2}{(1 - \epsilon_N)^2} \mu^2} \\
&= \frac{\sigma^2}{\frac{\epsilon_N^2}{(1 - \epsilon_N)} Np_N \mu}. \tag{C.14}
\end{aligned}$$

where the last inequality follows by Chebyshevs. Note that $\tau(p_N)$ is at least 1, we will apply this trivial lower bound along with *Eq. (C.14)* to bound the revenue,

$$\begin{aligned}
\mathcal{R}_{TB}(p_N) &= p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} + \zeta_N - \zeta_{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq N} \right] \\
&\leq p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} + \zeta_N \right) \mathcal{I}_{\tau(p_N) \leq N} \right] \\
&\leq p_N \zeta_N + p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} \right) \mathcal{I}_{\tau(p_N) \leq N} \right] \\
&= p_N \zeta_N + p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} \right) \mathcal{I}_{\tau(p_N) \in [N(1-\epsilon_N) \frac{p_N}{\mu}, N]} \right] \\
&\quad + p_N \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{N} \right) \mathcal{I}_{\tau(p_N) < N(1-\epsilon_N) \frac{p_N}{\mu}} \right] \tag{C.15}
\end{aligned}$$

$$\leq p_N \zeta_N + p_N \max \left\{ -\log \left((1 - \epsilon_N) \frac{p_N}{\mu} \right), 0 \right\} \tag{C.16}$$

$$+ p_N \left(-\log \left(\frac{1}{N} \right) \right) \frac{\sigma^2}{\frac{\epsilon_N^2}{(1 - \epsilon_N)} Np_N \mu} \tag{C.17}$$

$$= p_N \zeta_N + p_N \max \left\{ \log \frac{\mu}{(1 - \epsilon_N) p_N}, 0 \right\} + \frac{\sigma^2 \log N}{\frac{\epsilon_N^2}{(1 - \epsilon_N)} \mu N},$$

where Eq. (C.17) follows by the monotonicity of $\log(\cdot)$, and by applying Eq. (C.14) and $\tau \geq 1$. Now setting $\epsilon_N = N^{-\frac{1}{3}}$ and substituting gives,

$$\mathcal{R}_{TB}(p_N) \leq p_N \zeta_N + p_N \max\left\{\log \frac{\mu}{(1 - N^{-\frac{1}{3}})p_N}, 0\right\} + \frac{(1 - N^{-\frac{1}{3}})\sigma^2 \log N}{\mu N^{\frac{1}{3}}}. \quad (\text{C.18})$$

Maximizing Eq. (C.18) over p_N gives $\bar{p}_N := \mu / \exp(1 + \log(1 - N^{-\frac{1}{3}}) - \zeta_N)$. Combining Eq. (C.13) and Eq. (C.18), we have

$$\frac{\mu}{e} \log \left(\frac{1}{\frac{1}{e} + \frac{1 + \sigma^2}{N}} \right) \leq \mathcal{R}_{TB} \leq \bar{p}_N \left(\zeta_N - \log(1 - N^{-\frac{1}{3}}) \right) + \bar{p}_N \log \frac{\mu}{\bar{p}_N} + \frac{(1 - N^{-\frac{1}{3}})\sigma^2 \log N}{\mu N^{\frac{1}{3}}}. \quad (\text{C.19})$$

Taking limits of both sides of Eq. (C.19) completes the proof. \square

Omitted Proofs from Section 3.4

Proof of Corollary 3.4.1. We modify the random walk X_t in the proof of Theorem 3.3.1(a) into a stochastic process $\{X'_t, t \geq 0\}$. For $t \leq N$, let X'_t be the net utility of a random customer after opening t boxes. For $t > N$, simply let $X'_t - X'_{t-1} = \bar{\mu} - p$. For now X'_N has mean $N\bar{\mu}$ and variance $N\bar{\sigma}^2$. Also, the expectation of $X'_t - X'_{t-1}$ is $\bar{\mu}$ for any $t \geq 1$. Following the proof of Theorem 3(a), the Wald's equation and Chebyshev's inequality are still valid, so the result remains the same. \square \square

Proof of Lemma 3.4.1. Fix $k \in (0, 1)$ and let $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$, we will show that the normalized number of loot box purchases made by a customer under the pricing and allocation strategy (p, \mathbf{d}) , $\mathbb{E}[Q_{\mathbf{d}}^N(p)]$, tends to k as $N \rightarrow \infty$. Formally, for any $\epsilon > 0$, we will show there exists an integer M such that for all $N \geq M$, $k - \epsilon \leq \mathbb{E}[Q_{\mathbf{d}}^N(p)] \leq k + \epsilon$. For clarity we will separate the proof into two cases.

Lower Bound: $k - \epsilon \leq \mathbb{E}[Q_{\mathbf{d}}^N(p)]$

Given $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$, we first bound the probability that $Q_{\mathbf{d}}^N(p) < (1 - \epsilon)k$. Note since a customers valuation for the next loot box decreases after each purchase, this event is complementary to the event in which a customers valuation for the loot box dips below is less than p after they have opened $(1 - \epsilon)kN$ boxes. We will bound this event by applying Chebyshev's inequality, for which we will need estimates of both the mean and variance of customers valuation after opening $(1 - \epsilon)kN$ boxes. Let Z_i^m be an indicator random variable taking value 1 if item i from class m has not been revealed after $(1 - \epsilon)kN$ purchases, and 0 otherwise. When the class is clear from the context we will drop the superscript. Now, after each purchase the probability that item i in class m is obtained is $\frac{d_m}{\beta_m N}$, thus the expectation of Z_i^m is,

$$\mathbb{E}[Z_i^m] = \left(1 - \frac{d_m}{\beta_m N}\right)^{(1-\epsilon)kN}.$$

Let G_m denote the set of items in class m . For a random customer, the valuation of the next loot box after $(1 - \epsilon)kN$ purchases is given by $\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i^m$, and the expected valuation for a loot box after $(1 - \epsilon)kN$ purchases is,

$$\begin{aligned} \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i^m \right] &= \sum_{m=1}^M \sum_{i \in G_m} \mu_m \frac{d_m}{\beta_m N} \left(1 - \frac{d_m}{\beta_m N}\right)^{(1-\epsilon)kN} \\ &= \sum_{m=1}^M \mu_m d_m \left(1 - \frac{d_m}{\beta_m N}\right)^{(1-\epsilon)kN}. \end{aligned}$$

Moreover, for each class m the set of indicators $\{Z_i^m\}_{i=1}^{|G_m|}$ is negatively correlated since, for any $i, j \in G_m$, if Z_i^m is not revealed, Z_j^m is more likely to be revealed. Thus

the variance can be bounded by,

$$\begin{aligned}
& \text{Var} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right) \\
& \leq \sum_{m=1}^M \sum_{i \in G_m} \text{Var} \left(\frac{d_m}{\beta_m N} V_i Z_i \right) \quad (\{Z_i^m\} \text{ Neg. Coro.}) \\
& = \sum_{m=1}^M \sum_{i \in G_m} \left(\frac{d_m}{\beta_m N} \right)^2 (\mathbb{E}[V_i^2 Z_i^2] - (\mathbb{E}[V_i Z_i])^2) \\
& = \sum_{m=1}^M \sum_{i \in G_m} \left(\frac{d_m}{\beta_m N} \right)^2 (\mathbb{E}[V_i^2 Z_i] - (\mathbb{E}[V_i Z_i])^2) \quad (Z_i^m \in \{0, 1\}) \\
& = \sum_{m=1}^M \sum_{i \in G_m} \left(\frac{d_m}{\beta_m N} \right)^2 \left((\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN} - \mu_m^2 \left(1 - \frac{d_m}{\beta_m N} \right)^{2(1-\epsilon)kN} \right) \\
& \leq \sum_{m=1}^M \sum_{i \in G_m} \left(\frac{d_m}{\beta_m N} \right)^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN} \\
& = \sum_{m=1}^M \frac{d_m^2}{\beta_m N} (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN} \quad (\text{C.20})
\end{aligned}$$

Now applying Chebyshev Inequality to the event that less than , we have

$$\begin{aligned}
& \mathbb{P} (Q_{\mathbf{d}}^N(p) < (1 - \epsilon)k) \\
& = \mathbb{P} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i < p \right) \\
& = \mathbb{P} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] < p - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right) \\
& \leq \mathbb{P} \left(\left| \sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right| > \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] - p \right)
\end{aligned}$$

Now applying Chebyshev's inequality

$$\begin{aligned}
& \leq \frac{\text{Var} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right)}{\left(\mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] - p \right)^2} \\
& \leq \frac{\sum_{m=1}^M \frac{d_m^2}{\beta_m N} (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN}}{\left(\sum_{m=1}^M \mu_m d_m \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN} - p \right)^2} \quad (\text{Eq. (C.20)}) \\
& = \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN}}{\beta_m N \left(\sum_{m=1}^M \mu_m d_m \left(\left(1 - \frac{d_m}{\beta_m N} \right)^{(1-\epsilon)kN} - e^{-\frac{d_m}{\beta_m} k} \right) \right)^2} \quad (\text{C.21})
\end{aligned}$$

Taking the limit as N tends to infinity, the numerator of Eq. (C.21) approaches a constant, and the denominator goes to infinity. Thus for any $\epsilon > 0$, there exists M_1 such that for all $N > M_1$, $\mathbb{P}(Q_{\mathbf{d}}^N(p) < (1 - \epsilon/2k)k) \leq \epsilon/2k$. Applying Eq. (C.21) yields a lower bound of $\mathbb{E}[Q_{\mathbf{d}}^N(p)]$,

$$\begin{aligned}
\mathbb{E}[Q_{\mathbf{d}}^N(p)] &= \mathbb{E}\left[Q_{\mathbf{d}}^N(p) \mathcal{I}_{Q_{\mathbf{d}}^N(p) < (1-\epsilon/2k)k}\right] + \mathbb{E}\left[Q_{\mathbf{d}}^N(p) \mathcal{I}_{Q_{\mathbf{d}}^N(p) \geq (1-\epsilon/2k)k}\right] \\
&\geq 0 + \left(1 - \frac{\epsilon}{2k}\right) k \left(1 - \frac{\epsilon}{2k}\right) \\
&= k \left(1 - \frac{\epsilon}{2k}\right)^2 \\
&\geq k \left(1 - \frac{\epsilon}{k}\right) = k - \epsilon. \quad (\text{C.22})
\end{aligned}$$

Upper Bound: $\mathbb{E}[Q_{\mathbf{d}}^N(p)] \leq k + \epsilon$

As for the lower bound, we will first control the event that $Q_{\mathbf{d}}^N(p) > (1 + \epsilon)k$. Let $Z_i^m = 1$ now denote the event that after opening $(1 + \epsilon)kN$ loot boxes, item i in group m is still not revealed. As before we will omit the superscript when it is clear from context. Following the derivation of (C.21), we may bound the probability of

this event by,

$$\begin{aligned}
& \mathbb{P} \left(Q_{\mathbf{d}}^N(p) > (1 + \epsilon)k \right) \\
&= \mathbb{P} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i > p \right) \\
&= \mathbb{P} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] > p - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right) \\
&\leq \mathbb{P} \left(\left| \sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right| > p - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right)
\end{aligned}$$

Now applying Chebyshev's like before:

$$\begin{aligned}
&\leq \frac{\text{Var} \left(\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right)}{\left(p - \mathbb{E} \left[\sum_{m=1}^M \sum_{i \in G_m} \frac{d_m}{\beta_m N} V_i Z_i \right] \right)^2} \\
&\leq \frac{\sum_{m=1}^M \frac{d_m^2}{\beta_m N} (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN}}{\left(p - \sum_{m=1}^M \mu_m d_m \left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN} \right)^2} \quad (\text{Counterpart of } Eq. (C.20)) \\
&= \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN}}{\beta_m N \left(\sum_{m=1}^M \mu_m d_m \left(e^{-\frac{d_m}{\beta_m} k} - \left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN} \right) \right)^2}. \quad (C.23)
\end{aligned}$$

Now we will choose $\epsilon = -\log(1 - N^{-1/3})/k$. Substituting our choice of ϵ into the

denominator of Eq. (C.23) we may obtain a lower bound,

$$\begin{aligned}
& \beta_m N \left(\sum_{m=1}^M \mu_m d_m \left(e^{-\frac{d_m}{\beta_m} k} - \left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN} \right) \right)^2 \\
& \geq \beta_m \left(\sum_{m=1}^M \mu_m d_m N^{1/2} \left(e^{-\frac{d_m}{\beta_m} k} - e^{-\frac{d_m}{\beta_m} (1+\epsilon)kN} \right) \right)^2 \\
& = \beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} N^{1/2} \left(1 - \left(1 - N^{-1/3} \right)^{\frac{d_m}{\beta_m}} \right) \right)^2 \\
& \geq \beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} N^{1/2} \left(1 - \frac{1}{1 + \frac{d_m}{\beta_m} N^{-1/3}} \right) \right)^2 \quad \left(\left(1 - N^{-1/3} \right)^{\frac{d_m}{\beta_m}} \leq \frac{1}{1 + \frac{d_m}{\beta_m} N^{-1/3}} \right) \\
& = \beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} \frac{\frac{d_m}{\beta_m} N^{1/2}}{N^{1/3} + \frac{d_m}{\beta_m}} \right)^2.
\end{aligned}$$

where the second inequality follows from $\left(\left(1 - \frac{d_m}{\beta_m N} \right)^{(1+\epsilon)kN} \leq e^{-\frac{d_m}{\beta_m} (1+\epsilon)kN} \right)$. Plugging back into Eq. (C.23), the probability that customer purchases more than $(1 - \log(1 - N^{-1/3}))kN$ boxes is then bounded above by,

$$\mathbb{P} \left(Q_{\mathbf{d}}^N(p) > (1 - \log(1 - N^{-1/3}))k \right) \leq \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1 - \log(1 - N^{-1/3}))/k)kN}}{\beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} \frac{\frac{d_m}{\beta_m} N^{1/2}}{N^{1/3} + \frac{d_m}{\beta_m}} \right)^2}.$$

Finally, returning to $Q_{\mathbf{d}(p)}^N$, a trivial upper bound on $\mathbb{E}[Q_{\mathbf{d}}^N(p)]$ is the expected number of purchases necessary to obtain all N items. To compute this fix a class m , to collect all

the item in class m , on average the customer needs to open

$$\begin{aligned}
& \mathbb{E}[\# \text{ of purchases to collect all the items in class } m] \\
&= \mathbb{E} \left[\text{Geo}(d_m) + \text{Geo} \left(\frac{d_m(\beta_m N - 1)}{\beta_m N} \right) + \cdots + \text{Geo} \left(\frac{d_m}{\beta_m N} \right) \right] \\
&= \frac{\beta_m N}{d_m \beta_m N} + \frac{\beta_m N}{d_m(\beta_m N - 1)} + \cdots + \frac{\beta_m N}{d_m} \\
&= \frac{\beta_m N}{d_m} \left(\frac{1}{\beta_m N} + \frac{1}{\beta_m N - 1} + \cdots + 1 \right) \\
&\leq \frac{\beta_m N}{d_m} (\log(\beta_m N) + 1) \\
&\leq \frac{\beta_m N}{d_m} (\log N + 1).
\end{aligned}$$

The expected number of purchases required to collect all the items is bounded by the sum of the number of purchases to collect all the items in each class,

$$\begin{aligned}
\mathbb{E}[Q_{\mathbf{d}}^N(p)] &\leq \mathbb{E}[\# \text{ of purchases to collect all the items}] \\
&\leq \sum_{m=1}^M \mathbb{E}[\# \text{ of purchases to collect all the items in class } m] \\
&\leq \sum_{m=1}^M \frac{\beta_m N}{d_m} (\log N + 1).
\end{aligned}$$

Thus the expected number of purchases is upper bounded by $\mathbb{E}[Q_{\mathbf{d}}^N(p)] \leq \sum_{m=1}^M \frac{\beta_m N}{d_m} (\log N + 1)$ for any price p . Now we can build an upper bound as:

$$\begin{aligned}
\mathbb{E}[Q_{\mathbf{d}}^N(p)] &= \mathbb{E} \left[Q_{\mathbf{d}}^N(p) \mathcal{I}_{Q_{\mathbf{d}}^N(p) \leq (1+\epsilon)kN} \right] + \mathbb{E} \left[Q_{\mathbf{d}}^N(p) \mathcal{I}_{Q_{\mathbf{d}}^N(p) > (1+\epsilon)kN} \right] \\
&\leq \left(1 - \log(1 - N^{-1/3})/k \right) kN \\
&\quad + N \sum_{m=1}^M \frac{\beta_m}{d_m} (\log N + 1) \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N} \right)^{(1 - \log(1 - N^{-1/3})/k)kN}}{\beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k \frac{d_m}{N^{1/3} + \frac{d_m}{\beta_m}}} \right)^2}
\end{aligned}$$

Taking $N \rightarrow \infty$ on both sides, we have

$$\begin{aligned}
\lim_{N \rightarrow \infty} \frac{\mathbb{E}[Q_{\mathbf{d}}^N(p)]}{N} &\leq \left(1 - \log(1 - N^{-1/3})/k\right) k \tag{C.24} \\
&+ \sum_{m=1}^M \frac{\beta_m}{d_m} (\log N + 1) \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) \left(1 - \frac{d_m}{\beta_m N}\right)^{(1 - \log(1 - N^{-1/3})/k)kN}}{\beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} \frac{\frac{d_m}{\beta_m} N^{1/2}}{N^{1/3} + \frac{d_m}{\beta_m}}\right)^2} \\
&= k + \lim_{N \rightarrow \infty} \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) e^{-\frac{d_m}{\beta_m} k} \sum_{m=1}^M \frac{\beta_m}{d_m} (\log N + 1)}{\beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} \frac{\frac{d_m}{\beta_m} N^{1/2}}{N^{1/3} + \frac{d_m}{\beta_m}}\right)^2} \\
&= k + \lim_{N \rightarrow \infty} \sum_{m=1}^M \frac{d_m^2 (\mu_m^2 + \sigma_m^2) e^{-\frac{d_m}{\beta_m} k} \sum_{m=1}^M \frac{\beta_m}{d_m} \left(1 + \frac{1}{\log N}\right)}{\beta_m \left(\sum_{m=1}^M \mu_m d_m e^{-\frac{d_m}{\beta_m} k} \frac{\frac{d_m}{\beta_m} N^{1/2}}{(\log N)^{1/2} (N^{1/3} + \frac{d_m}{\beta_m})}\right)^2} \\
&= k.
\end{aligned}$$

Thus for any $\epsilon > 0$, there exists M_2 such that for $N > M_2$, the RHS of (C.25) $< k + \epsilon$. Taking the maximum of M_1 and M_2 , and combining Eqs. (C.22) and (C.25), we know that for any $N > \max(M_1, M_2)$,

$$\left| \frac{\mathbb{E}[Q_{\mathbf{d}}^N(p)]}{N} - k \right| < \epsilon \implies \lim_{N \rightarrow \infty} \frac{\mathbb{E}[Q_{\mathbf{d}}^N(p)]}{N} = k.$$

□

Proof of Theorem 3.4.1. Consider a finite group allocation $\mathbf{d} = (d_1, \dots, d_M)$. For $p > \sum_{m=1}^M d_m \mu_m$, by the law of large numbers, the normalized selling volume will tend to 0. Now suppose $p \leq \sum_{m=1}^M d_m \mu_m$. Note $\sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$ equals $\bar{\mu}$ when $k = 0$, and decreases monotonically to 0 as $k \rightarrow \infty$. Thus for any such p , there exist a unique positive k such that $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$. Recall by Lemma 3.4.1, if $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k}$, $k > 0$, then

$$\lim_{N \rightarrow \infty} \mathbb{E}(Q_{\mathbf{d}}^N(p)) / N = k.$$

Using this identity we can write the limiting revenue function in terms of k , i.e.,

$$\lim_{N \rightarrow \infty} \mathcal{R}(p, \mathbf{d}) = \lim_{N \rightarrow \infty} p \cdot \mathbb{E}(Q_{\mathbf{d}}^N(p)) / N = k \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k} := \sum_{m=1}^M G_m(k).$$

Consider the m^{th} term of the revenue function, $G_m(k) = \mu_m d_m k e^{-\frac{d_m}{\beta_m} k}$. This function obtains its maximum at $k = \beta_m / d_m$, and the maximum value is $\beta_m \mu_m / e$, which is independent from the value of d_m . Hence, the revenue is bounded by $\sum_{m=1}^M \beta_m \mu_m / e = \bar{\mu} / e$. For any \mathbf{d} , we can reach the the upper bound $\bar{\mu} / e$ only if every component function reaches the maximum simultaneously, i.e., $k = \beta_m / d_m$ for all m . Since $\sum_{m=1}^M \beta_m = \sum_{m=1}^M d_m = 1$, the only possible allocation is $d_m = \beta_m$, which is proportional allocation. In this case, $k = 1$, and the corresponding price is $p = \sum_{m=1}^M d_m \mu_m e^{-\frac{d_m}{\beta_m} k} = \bar{\mu} / e$. Hence the optimal solution is $p = \bar{\mu} / e$ with proportional allocation, and the optimal revenue is $\bar{\mu} / e$. \square

Proof of Proposition 3.4.1. Let p be the price separate selling uses. Now consider a loot box strategy (either unique box or traditional box) with salvage cost p and price p . The customer will purchase indefinitely, keeping all the items which they value at p or greater, and returning the unwanted items for a full refund. Hence, such a loot box induces the same revenue as separate selling, which implies that

$$\max_c \mathcal{R}_{UB}^c, \max_c \mathcal{R}_{TB}^c \geq \mathcal{R}_{SS}.$$

\square

Proof of Theorem 3.4.2. We will consider the two cases separately.

(a) Let $V_i^c = \max\{V_i, c\}$ and F_c be the valuation of item i and the distribution of customer valuations under the a salvage system with cost c . Let $\tilde{\eta}$ be the mean of

F_c . By Theorem 3.3.1(a), as the number of items $N \rightarrow \infty$, the optimal price tends to $\tilde{\eta}$ and the expected proportion of items obtained tends to 1. Since almost all items are obtained in expectation, the proportion of items salvaged is $F(c) = \gamma$ w.p. 1. Thus the normalized cost of salvages by the customer is $\lim_{N \rightarrow \infty} \frac{\# \text{ Items Salvaged}}{N} c = \gamma c$. Together, the normalized revenue is then $\tilde{\eta} - \gamma c$. Noting $\tilde{\eta}$ can be rewritten as $E[\max\{V, c\}] = \gamma c + (1 - \gamma)\mathbb{E}[V|V > c] = \gamma c + (1 - \gamma)\eta$ and plugging in gives the result.

(b) Following Theorem 3.3.1(b), we will consider the a modified sum of valuations random walk for customers of traditional box with salvage cost c . First, consider partition the items into two sets based on the items that will be salvaged, $S_L = \{i|V_i \leq c\}$, and items that will not be salvaged $S_U = \{i|V_i > c\}$. Index the items in S_U from 1 to $|S_U|$ in reverse order in which they will be allocated $\{V_i^U\}_{i=1}^{|S_U|}$. Then the sum of valuations random walk is $\{Y_j\}_{j=0}^{|S_U|}$, where $Y_j = c(|S_L| + |S_U| - j) + \sum_{i=1}^j V_i^U$ and $Y_0 = cN$. For a random customer, $Y_{|S_U|}/N$ is their expected valuation for the first traditional box. Each time the customer receives a new item from S_U , they move back one index on the random walk from Y_{j+1} to Y_j . Then $\{Y_i\}$ is strictly increasing, since every time the customer receives a new item from $|S_U|$, their total valuation for the remaining items decreases.

Fix an N and price p_N . Then a customer purchases until the first time t (counting down from $|S_U|$) such that $\frac{Y_t}{N} < p_N \iff Y_t < Np_N$. Let $\tau(p_N)$ be the first passage time of $\{Y_i\}$ to Np_N i.e., $\tau(p_N) := \min\{t : Y_t \geq Np_N\}$. Note that since Y_i is monotonic, the random walk crosses Np_N exactly once. Suppose $\tau(p_N) \leq |S_U|$, then at time $\tau(p_N)$ the number of distinct items a customer will own and keep is $|S_U| - \tau(p_N) - 1$, and the number of traditional loot boxes they will have purchased is the sum of $|S_U| - \tau(p_N) - 1$ independent geometric random variables, $\text{Geo}(\frac{|S_U|}{N}) + \text{Geo}(\frac{|S_U|-1}{N}) + \dots + \text{Geo}(\frac{\tau(p_N)}{N})$. Then the revenue will be $\frac{(p_N - c) \times \# \text{ Purchases}}{N} +$

$\frac{c \times \# \text{ Items in } |S_U| \text{ Obtained}}{N}$. That is,

$$\begin{aligned} \mathcal{R}_{TB}^c(p_N) &= (p_N - c) \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{|S_U|} + \zeta_{|S_U|} - \zeta_{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq |S_U|} \right] \\ &\quad + c \mathbb{E} \left[\frac{|S_U| - \tau(p_N) - 1}{N} \mathcal{I}_{\tau(p_N) \leq |S_U|} \right]. \end{aligned} \quad (\text{C.25})$$

Now we will bound $\mathbb{E}[\tau(p_N)]$. Recall $\eta = E[V|V > c]$ is the mean valuation of an item in S_U . Since $\tau(p_N)$ is the passage time, by Wald's equation, $\mathbb{E}[Y_{\tau(p_N)}] = cN + \mathbb{E}[\tau(p_N)](\eta - c)$, we have

$$\mathbb{E}[\tau(p_N)] = \frac{\mathbb{E}[Y_{\tau(p_N)}] - cN}{\eta - c} \in \left[N \frac{p_N - c}{\eta - c}, N \frac{p_N - c}{\eta - c} + O\left(\frac{\sigma^2}{\mu^2}\right) \right].$$

Note $|S_U|$ is binomial with mean $(1 - \gamma)N$. Conditioning on the event that $S_U \in [(1 - \gamma)N - \delta, (1 - \gamma)N + \delta]$ we have a lower bound on the revenue,

$$\begin{aligned} \mathcal{R}_{TB}^c(N, p_N) &\geq (p_N - c) \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{|S_U| - \delta} + \zeta_{|S_U|} - \zeta_{\tau(p_N)} \right) \mathcal{I}_{\tau(p_N) \leq N, |S_U - (1 - \gamma)N| \leq \delta} \right] \\ &\quad + c \mathbb{E} \left[\frac{|S_U| - \tau(p_N) - 1}{N} \mathcal{I}_{\tau(p_N) \leq |S_U|} \right] \\ &\geq ((p_N - c) \mathbb{E} \left[\left(-\log \frac{\tau(p_N)}{|S_U| - \delta} \right) \right] \\ &\quad + c \mathbb{E} \left[\frac{|S_U| - \tau(p_N) - 1}{N} \right]) \Pr(|S_U - (1 - \gamma)N| \leq \delta) \\ &\geq -(p_N - c) \log \mathbb{E} \left[\frac{\tau(p_N)}{|S_U| - \delta} \right] \\ &\quad + c \mathbb{E} \left[\frac{|S_U| - \tau(p_N) - 1}{N} \right] \Pr(|S_U - (1 - \gamma)N| \leq \delta). \end{aligned}$$

where the final inequality follows from Jensens. Plugging in our upper bound for $\mathbb{E}[\tau(p_N)]$, applying straight forward concentrations on S_U , and taking the limit yields,

$$\lim_{N \rightarrow \infty} \mathcal{R}_{TB}^c(p, N) \geq (p - c) \log \left(\frac{(1 - \gamma)(\eta - c)}{p - c} \right) + c \left(1 - \gamma - \frac{p - c}{\eta - c} \right). \quad (\text{C.26})$$

Maximizing over the price yields $p = c + e^{-\frac{\eta}{\eta-c}}(1-\gamma)(\eta-c)$. Plugging in p gives our desired revenue $(1-\gamma)(\eta-c) \left(\frac{c}{\eta-c} + e^{-\frac{\eta}{\eta-c}} \right)$. An upper bound on the revenue follows by an argument analogous Theorem 3.3.1(b), we omit the details for brevity. \square